Comments on “Threshold values and management options for nutrients in a catchment of a temperate estuary with poor ecological status” of Hinsby et al. (2012)

1.) General remarks and suggestions

The authors have provided a very interesting mauscript, which deals with an integrated identification of nitrogen and phosphorous threshold values for both surface and subsurface runoff to a Danish estuary. The approach presented is mainly based on models and data that are nationally and regionally available in Denmark. Such kind of combined calculation and modelling concept is highly needed within the context of the second generation establishment of River Basin management Plans to fulfil the requirements of the Water framework Directive. Moreover the paper clearly can be characterized as trans-disciplinary.

However the submitted paper version suffers form of some shortcomings which should be addressed by the authors prior to publication.

1. Some of the arguments are weak or need some references or further explanations. This applies particularly the modelling parts, where uncertainty and robustness estimations are missing.

2. There are parts of the discussion on the ecological status and acceptable loads in chapter 4.2 and 5.3 that should moved to the “methods chapter” as they include information that are absolutely vital to understand the discussion on specific thresholds. This applies particularly the derivation of the 14 µg DIN L^{-1} and 6.2 µg DIP L^{-1} nutrient limitation values

3. It is understandable that the concept presented involves various disciplines, data and modelling concepts. However, the authors might consider condensing some of the discussions by referring to easily accessible articles.

2.) Specific comments

Specific comments, covering both, questions, remarks and suggestions for corrections / improvements are provided in the attached text version of the paper provided.

Threshold values and management options for nutrients in a catchment of a temperate estuary with poor ecological status

K. Hinsby[1], S. Markager[2], B. Kronvang[3], J. Windolf [3], T. Sonnenborg[1] and L. Thorling[4].

(1){Dept. Hydrology, Geological Survey of Denmark and Greenland, Copenhagen, Denmark}

(2){Dept. Bioscience, Aarhus University, Roskilde, Denmark}

(3){Dept. Bioscience, Aarhus University, Silkeborg, Denmark}

(4){Dept. Groundwater and Quaternary Geology mapping, Geological Survey of Denmark and Greenland, Aarhus, Denmark}

Correspondence to: K. Hinsby (khi@geus.dk)

Abstract

Intensive farming has severe impacts on the chemical status of groundwater and streams and consequently on the ecological status of dependent ecosystems. Eutrophication is a widespread problem in lakes and marine waters. Common problems are hypoxia, algal blooms and fish kills, ~~and~~ loss of water clarity, underwater vegetation, biodiversity, and recreational value. In this paper we evaluate the nitrogen (N) and phosphorus (P) chemistry of groundwater and surface water in a coastal catchment, the loadings and sources of N and P and their effect on the ecological status of an estuary. We calculate the necessary reductions in N and P loadings to the estuary for obtaining a good ecological status, which we define based on the number of days with N and P limitation, and the equivalent stream and groundwater threshold values assuming two different management options. The calculations are performed by the combined use of empirical models and a physically based 3D integrated hydrological model of the whole catchment. The assessment of the ecological status indicates that the N and P loads to the investigated estuary should be reduced to levels corresponding to 52 and 56 % of the current loads.~~by a factor of 0.52 and 0.56~~, respectively, to restore good ecological status. Model estimates show that threshold total N concentrations should be in the range of 2.9 to 3.1 mg L^{-1} in inlet freshwater to Horsens Estuary and 6.0 to 9.3 mg L^{-1} in shallow aerobic groundwater (~ 27 – 41 mg L^{-1} of nitrate), depending on the management measures implemented in the catchment. The situation for total P is more complex but data indicate that groundwater threshold values are not needed. The inlet freshwater threshold value for total P to Horsens Estuary for the selected management options is 0.084 mg L^{-1}. Regional climate models project increasing winter precipitation and runoff in the investigated region resulting in increasing runoff and nutrient loads to coastal waters if present land use and farming practices continue. Hence, lower threshold values are required in the future to ensure good status of all water bodies and ecosystems.

1 Introduction

Nutrient emissions from anthropogenic sources may have severe impacts on the water environment and cause significant problems with the chemical status of water resources and the ecological status of terrestrial, freshwater, and marine ecosystems globally (Vitousek et al., 1997; Tilman et al., 2001; Galloway et al., 2004; Diaz and Rosenberg, 2008; Rockstrøm et al., 2009). Rockstrøm et al. (2009) identify the human impact on the biogeochemical cycle of nitrogen as one of the currently most severe environmental problems globally and recommend~~s~~ that the human fixation of nitrogen and emissions of reactive nitrogen species are reduced to 25% of the present levels. Hence, there is a strong and increasing need to regulate and reduce nutrient loadings, particularly in areas with intensive farming, in order to protect water resources and ecosystems (Tilman et al. 2001; Rockstrøm et al., 2009).

The European Groundwater Directive (EU, 2006) stipulates that the European Union (EU) member states have to derive groundwater threshold values for all relevant contaminants in all groundwater bodies that may put associated ecosystems at risk. These risks include harmful algal blooms, hypoxia, and loss of biodiversity and underwater vegetation in aquatic ecosystems (Cloern 2001; Conley et al., 2002; Hinsby et al., 2008). Groundwater threshold values are concentrations which should not be exceeded in order to assure good chemical and ecological status of groundwater associated or dependent ecosystems. If the threshold value for a given pollutant is exceeded the groundwater body is classified as having poor chemical status according to EU directives (EU, 2000, 2006). Presently, the EU directives do not require a similar derivation of stream threshold values. However, we recommend that stream and groundwater threshold values are derived together, as stream threshold values can be calculated directly from estimated maximum nutrient loads to lakes and marine areas.

Comment [d1]: Why? The reader migh assume that stream threshold values relate to the ecological status of the streams and are not necessarily linked to the maximum loads to lakes and marine areas. cf.: J.A. Cam Argo, A. Alonso, A. Salamanca (2005): Nitrate toxicity to aquatic animals: a review with new data for freshwater invertebrates, Chemosphere, 58 (9) (2005), pp. 1255–1267.
Hence the authors definition of stream threshold values needs with this respect to be limited to the importance of stream water chemistry for the respective marine areas.

An integrated assessment of threshold values for groundwater based on targets for protection of associated or dependent ecosystems is an interdisciplinary challenge that needs contributions from disciplines like marine and freshwater ecology, hydrology, hydrogeology, and hydrochemistry, as well as data for all water bodies in the investigated hydrological system. To the authors' knowledge, this is the first interdisciplinary study that estimates groundwater threshold values based on targets for the ecological status of a marine ecosystem. In this paper we: 1) calculate total land based nitrogen and phosphorus loads 2) estimate maximum acceptable nitrogen and phosphorus loads to the estuary in order to ensure a good ecological status of the estuary 3) derive the equivalent nitrogen and phosphorus groundwater and stream threshold values for protection of the estuary and 4) assess the present chemical status of groundwater in the catchment to Horsens estuary relative to the derived groundwater threshold values.

Our aim is to provide and demonstrate a methodology for derivation of threshold values and integrated assessment of nutrient transport across hydrological systems, from groundwater to estuaries, using Horsens estuary and its catchment as an example, and establish the knowledge base and system understanding to assess the impacts of projected climate change on the evolution of the quantitative, chemical and ecological status in the investigated catchment in a companion paper.

2 Study area

The catchment: The area of investigation is a 518 km^2 Danish coastal catchment including the small islands in the estuary (Fig. 1). The catchment consists of two major gauged sub-catchments with gauging stations just upstream the two major lakes in the area, discharging about 70% of the freshwater from the total catchment through the two lakes into the inner western part of the estuary. A number of smaller ungauged sub-catchments are discharging to the estuary via a number of small streams on both sites of the estuary (Fig. 1). The dominant land use is agriculture (76%). The remaining areas are forested (10%), or lakes, wetlands and meadows (5%) *(*BLST, 2010*)*. The population in the area is about 110.000 (136 inhabitants per km^2), of which ~~and about~~ 73% live~~s~~ in municipalities with sewer systems. The animal production is dominated by pigs (69%) and cows (26%), and the area currently contains 0.79 livestock animal units (AU) per hectare agricultural soil *(*BLST, 2010*)*.

The geology and topography of the area was developed by glacial processes during the last glaciation (Weichselian / Wisconsinian). The deposits are mainly clay tills and outwash sands constituting the main aquitards and aquifers, although some glaciolacustrine clay layers also exist. A conceptual model of the geological and hydrological setting in the catchment with indication of type of available data, nutrient sources and transport~~-~~, is shown in Fig. 2.

There are five lakes located in the catchment (total surface area: 2.43 km^2), around 1700 ponds (total surface area: 2.21 km^2), and the catchment is drained by 595 km of streams of which 78% are less than 2 m wide. The mean precipitation for the agro-hydrological years 2000 to 2005, the period we model in this study, was 695 mm yr^{-1} and the corresponding total discharge from the catchment to the estuary was 299 mm yr^{-1}.

The estuary: The Horsens estuary is a shallow estuary with a mean depth of 2.9 m and a surface area of 77.5 km^2 (Stedmon et al 2006; Markager et al. 2011). Tidal range is low and mixing is mainly wind driven (Gustafsson and Bendtsen, 2007). The estuary is connected to the Belt Sea and the Baltic Sea transitions zone through a deep (16 m) channel and is generally well mixed with salinities from 12 to 26‰, which is comparable to the salinity in the Belt Sea. Despite the well mixed conditions, results from a 3D-ecological modeling study (Timmermann et al. 2010) show that the ecological conditions in the estuary are mainly governed by local nutrient inputs with the nutrient concentrations in the adjacent sea only playing a minor role. The nutrient concentrations in the estuary are typical for Danish estuaries and at the same order of magnitude as ~~comparable to~~ estuaries in the U.S. such as the Patuxent river estuary and Chesapeake Bay (Boynton and Kemp, 2008).

3 Material and Methods

3.1 Monitoring in the Horsens Fjord catchment and estuary

The first Danish Action Plan for the Aquatic Environment was adopted in 1987 and the resulting monitoring program has been in place since 1989. Hence more than 20 years of

Comment [dimul2]: Some of the legends and some texts are difficult to read in the pdf-version submitted, maybe due to problems when converting the original file to pdf format?
The authors might also reconsider the figure, as it indicates fx. that B=A+F, i.e. modelled stream loading equals measured concentration, runoff in streams and N retention. Rather the figure should help the reader to understand, when and how model results are used and how and when measurements contribute to the the approach

Comment [d3]: The authors might specify first and final month of a hydrological year monitoring data are presently available for all major water bodies (Kronvang et al., 2008; Hinsby and Jørgensen., 2009; Markager et al., 2010; Hansen et al.; 2011).

In this study we use data from this program collected in the investigated catchment and data from a small agricultural research and monitoring site a few kilometers outside the catchment, with intensive monitoring of tile drainage water and upper groundwater (1 to 5 meter below ground surface).

Discharge and nutrient concentrations are measured in the Bygholm and Hansted streams at the two gauging stations (Fig. 1) covering the discharge and loadings from 56% of the catchment area. Water sampling in streams was normally conducted every second week and analyzed for total nitrogen, nitrate-nitrite-N, ammonium-N, total P, and dissolved orthophosphate. Instantaneous discharge (Q) was measured 12 to 20 times per year using a low friction propeller, and daily discharge values were calculated using relationships between Q and continuously measured fluctuations in water level (H) in the streams.

Comment [d4]: Are monthly to bi-monthly Q measurements sufficient to reliably derive daily values based on a Q/H relationship? Can you provide figures that may help estimating uncertainty?

Monitoring in the estuary was initiated in 1980 and systematically collected data exist~~s~~ from 1985 to 2007. Monitored parameters included profiles of salinity, temperature, chlorophyll fluorescence, and light attenuation from CTD cast, as well as nutrient and chlorophyll concentrations from discrete water samples at two depths. Biomass measurements of underwater vegetation and the benthic invertebrates were performed together with enumeration of phytoplankton. The only rate measurement ~~was~~ has been performed for phytoplankton primary production. The sampling frequency varied from 12 to 46 times per year. Generally, sampling and analytical procedures follow Danish and European standards and directives i.e. most recently the requirements described in (EU, 2009). Selected data from the monitoring programs are shown in Table 1 and 2.

3.2 Data analysis and development of the conceptual model

For the derivation of stream and groundwater threshold values we apply a stepwise approach (Fig. 2). Firstly, the current N and P loadings to the estuary were estimated. Based on these values and empirical models for the relationships between loadings and nutrient concentrations acceptable N and P loadings to the estuary were estimated. Secondly, two scenarios were constructed for achieving these values for annual nutrient loading. Finally, these annual loadings were converted to groundwater and stream threshold values using a catchment model and monitoring data for N and monitoring data and expert judgment for P (Fig. 2).

3.2.1 Calculation of freshwater discharge, nutrient sources and loads

Monthly freshwater discharge and transport of nutrients are calculated using a linear interpolation method (Kronvang & Bruun, 1996) by multiplying daily nutrient concentrations with mean daily discharge calculated from stage-discharge relationships, developed for each of the the two gauging stations situated in the main stream inlets (Fig. 1). Land based monthly nutrient loadings and freshwater discharge from the entire catchment to the Horsens estuary for the period 1984 to 2009 have been estimated utilizing data from the two gauged stations, and adding modeled monthly freshwater discharge and nutrient loadings from the ungauged part of the catchment by using the DK-QN model complex according to Windolf et al. (2011) (Fig.2.). The DK-QN model is a combination of empirical nutrient loss models and the physically distributed and integrated hydrological DK-model (Henriksen et al., 2003). The modeled freshwater discharge for the ungauged catchment is derived from the DK-model (the Danish National Water Resource Model), which is based on the integrated hydrological modeling system MIKE SHE and in the second generation of the model, established for a grid size of 500 m x 500 m for the entire Denmark. In the present study the grid size has been

Comment [d5]: How often are nutrient concentration measurements performed, if not daily? Are these values estimated from measurements performed between 12 and 46 times a year? If yes, can you provide figures on uncertainty?

Comment [d6]: Please explains the assumptions that justify the application of the modelling concept to ungauged catchments?

Comment [d7]: MIKE SHE and in the second generation of the model? MIKE SHE comprises the codes and the DK-model can be regarded as a specific model setup for the entire Danish land mass. Wha do you mean by second generation?

refined further and reduced to 250 m x 250 m. Monthly nitrogen loadings were also modeled for the two gauged catchment thus allowing a validation of the applied DK-QN model complex against measured nitrogen concentrations at the two gauged stations. Moreover, the nitrate leaching from the root zone (upper 1 m) was calculated for the entire catchment to the Horsens estuary using the Danish empirical NLES leaching model, which performed well in a large inter-comparison with seven other well known nutrient models (Kronvang et al., 2009b).

Comment [d8]: The authors might be more specific on the models performance, particularly as the study mentioned has been performed at what the authors call large scale

For phosphorus, monthly loadings have been provided from the regional environmental authorities. The total loadings were apportioned to sources according to Table 3. The discharges from point sources were measured at the outlet (IP's, WWTP's, and FF's), or calculated based on treatment facilities and number of houses in each subcatchment, and experience data for production of nutrients and reduction efficiency of treatment (SD). The atmospheric deposition of nitrogen to fresh surface waters (A_{fresh}), and the surface area of the Horsens estuary (A_{marin}), was calculated based on national models for transportation and deposition (http://www.air.dmu.dk). Natural background losses of total nitrogen (NB) were estimated as flow-weighted concentrations from sampling in streams draining uncultivated catchments. The gross nutrient emission to and load in streams (L_s), was calculated by the established model and includes the loads described by equation (1):

Comment [d9]: Is it possible to assess the applicability of those data to the specifi area / study?

Comment [d10]: Based on what references / assumptions?

Comment [d11]: What is the apportionment based upon?

Comment [d12]: Please introduce definitions first before providing abbreviations

$$L_s = L_{agri} + L_{nb} + L_{ps} + L_{af} - R_{slw} \qquad (1)$$

hence the agricultural share of the gross nutrient emissions (L_{agri}) can be calculated by equation (2):

$$L_{agri} = L_s - L_{ps} - L_{nb} - L_{af} + R_{slw} \qquad (2)$$

where

L_s is the average ~~river-borne~~ loading of nutrients to the Horsens Estuary estimated from diffuse sources and according to the combined use of monitoring and modeling data; L_{ps} is the nutrient loads from point sources; L_{nb} is the natural background loads of nutrients from non-agricultural areas; L_{af} is the direct atmospheric deposition on surface freshwater; and R_{slw} is the retention of nutrients in the catchment after their emission to surface waters.

Comment [d13]: Consider moving theses explanations right behind eq. 1

3.2.2 Estimating maximum acceptable nutrient loads to Horsens Estuary

The estimation of maximum acceptable loads to Horsens Estuary was based on empirical models for relationships between N and P loadings and resulting N and P concentrations (effects) in the estuary (Fig. 4). The specific effects (y-variable) evaluated were annual mean concentration of total N and P mean concentrations of DIN from May through October and for DIP from March through July (Table 4). The periods for DIN and DIP correspond approximately to the periods were N or P limitation of phytoplankton occur in the estuary (data not shown). The empirical models were developed with an iterative multiple linear regression procedure working on standardized time series (zero mean and a standard deviation equal to one). The explanatory variables (x-variables) were N and P loads, water temperature, wind speed (cubed daily mean values), surface irradiance, salinity (used as a proxy for water exchange with the adjacent Belt Sea) and the North Atlantic Oscillation Index (NAO, http://www.cru.uea.ac.uk/~timo/projpages/nao_update). These variables represent~~s~~ the major external factors governing the conditions in the estuary, i.e. nutrient loadings, climatic forcing and water exchange. Each explanatory variable was calculated as mean values for eleven

Formatted: English (U.K.)

Comment [d14]: Lower figures: inlet o outlet (contradiction between figure text and legend text)

Comment [d15]: Please define DIN (as you do in figure 6 and on later in the text) and DIP

Comment [d16]: Standardized time series of what, the concentrations or the explanatory variables or both?

different time periods prior to and/or including the period for the response variable in order to allow for time lag between e.g. loads and resulting effects in the estuary. The eleven periods were: period 1 to 5) the periods for the response variable including 0, 1, 2, 4 and 8 months before and period 6) all months back to January in the previous year. Period 7 to 11) were periods ending when the response period started and starting 1, 2, 4 and 8 month before, and January in the previous year. This method gave 7 x 11 potential explanatory variables. A forward selection procedure adopted from *Broadhurst et al.* (1997)) was used to select the explanatory parameters (between two and five) providing the best model fit. A jack-knifing procedure was used to test all variables and all combinations of years and the best explanatory variables were chosen based on root mean square error of cross validation (RMSECV). Nitrogen and phosphorous loadings were always chosen as the first variable for their respective concentrations and only one variable for each class of explanatory variable was chosen, but otherwise the selection procedure for explanatory variables was based on RMSECV. The procedure stopped when further explanatory variables did not improve the model based on RMSECV (two to four explanatory variable were used). Time series from 1985 to 2006 were used i.e. 22 years, however, the last four years where not used in the parameter selection procedure but retained for validation. After validation of the explanatory parameter selection, a final estimation of the regression coefficients was done including all 22 years. The final results from the models are coefficients for the effects of changes in response variables per unit change in loadings (% change in response variable /% change in loading), adjusted for effects of inter annual variability in climatic conditions. These coefficients were subsequently used to estimate the values for response variables under reduced loadings assuming average climatic conditions, i.e. the final model equations were used as scenarios where N and P loads varied, but with climatic variables set to their average value in the data set. Finally, the maximum acceptable loads to the estuary were estimated using the calculated

Comment [d17]: Check against reference

Comment [d18]: Why did you decide t limit the final choice to max 5 parameters out of 77?

Comment [d19]: Was there any difference between the original, the validation and the final model?

Formatted: English (U.K.)

relationships between DIN and DIP mean concentrations, and the percentage of days with N and P limitation in the estuary (see section 4.3 and Fig. 8 for estimation of N and P limitation and 5.2 for a discussion of good ecological status).

Comment [d20]: It is understandable to consider "percentage of days with nutrients limitations in the estuary", but I am missing discussion of ecologically defined threshold values for nutrient limitations mentioned here. Given the title of the sub chapter any derivation of maximum concentrations beyond which negative affects of nutrient inputs to the estuary are enhanced should b explained and discussed here, rather than ir chapter 4.3 and 52.)

3.3 Scenarios of mitigation measures

The reduction targets for nutrient loadings calculated for the Horsens estuary can be accomplished by utilizing different mitigation measures in the catchment, and it is important to note that the actual selection of applied mitigation measures will affect the calculated groundwater threshold value for total N. The reason for this is that the chosen measures may include and take advantage of subsurface reduction (retention) processes to various degrees. Generally, the most strict groundwater threshold values would be established if subsurface retention is not increased and the reduction in nutrient loading is solely to be obtained by reducing the nutrients leaching from agricultural soils. Groundwater threshold values can be allowed to be higher if in addition other measures such as introduction of uncultivated buffer zones, restoration of wetlands along streams and reduction in other significant nutrient sources were applied to help reducing the nutrient loading to streams and ultimately the estuary. We have evaluated two possible scenarios to illustrate how the choice of mitigation measure will influence the derived groundwater threshold value for total N:

Scenario 1. Assumes that the entire reduction target for N and P is directed against the diffuse sources in the catchment, i.e. losses from fields. This scenario results in the lowest (most strict) groundwater threshold values.

Scenario 2. Measures are imposed on point sources, direct atmospheric deposition (through lower emission of ammonia from agriculture/manure) and diffuse sources. Furthermore, construction/restoration of wetlands and uncultivated buffer zones along streams were included for additional removal of nutrients. As this scenario utilizes further nitrogen reduction from other sources it allows higher threshold values in aerobic groundwater.

3.4 Derivation of stream threshold values

In contrast to groundwater threshold values, stream threshold values are not sensitive to the selected nutrient management option in the investigated catchment. The flow-weighted stream concentrations simply has to be reduced by the same relative amount as required for the estuary as the stream input constitute approximately 90% of the total nutrient input to the estuary (Table 3).

Comment [d21]: Please provide an explanation.

To estimate the current total N loading from streams to the estuary and the required threshold values, we have applied an empirical model for estimating monthly flow-weighted total nitrogen concentrations in freshwater discharge to minor streams. The model was developed based on nitrogen data for 83 small agricultural catchments without lakes and wetlands and data for the period 1990 to 2009 using an approach described by Kronvang et al. (1995), Andersen et al. (2005), and Windolf et *al*. (2011). The retention of total nitrogen in streams, lakes and wetlands was calculated utilizing different models and expert judgments as described in Windolf et al. (2011) and Kronvang et *al.* (2005). The modeling complex allowed a model estimation of gross and net stream flow-weighted concentrations taking into consideration the nutrient retention in the 5 larger lakes situated in the catchment of which the 2 largest are situated downstream the two monitoring stations just before river water enters the estuary (Fig. 1). Net inlet freshwater nitrogen threshold values to Horsens Estuary were

Comment [d22]: Are any subcatchments of the Horsens estuary catchment among the 83 catchments or is there any other chance to judge, how the study sub catchments are covered by the model e.g. as to the ranges of the parameter sets?

calculated utilizing this model complex for the two scenarios. The threshold values for total phosphorus were calculated as net flow-weighted concentrations.

Comment [d23]: The sections explains how concentrations are calculated from loadings, and it is claimed that concentrations have to be reduced to 90 % of the current level. However, given the title of the section, I am missing a description, how exactly threshold values are identified.

3.5 Derivation of groundwater threshold values

Groundwater threshold values depends on the application of possible mitigation measures as described in section 3.3. The threshold value has to be calculated for aerobic groundwater as the major nitrogen species in groundwater, nitrate, are reduced to unreactive N_2 at the redox boundary.

Dissolved inorganic nitrogen (DIN = NO_2-N + NO_3-N + NH_4-N) in anaerobic groundwater in the investigated catchment is primarily present as ammonia at concentrations that are generally 1-2 orders of magnitude lower than the DIN concentrations in aerobic groundwater, where nitrate is the dominant nitrogen species (Table 1). Hence the major part of the total N load to streams is generally nitrate originating from shallow aerobic groundwater that discharge either directly or via drainage ditches or tiles to the stream (Fig. 2). As there are only few monitoring wells in aerobic groundwater in the investigated catchment the leaching of nitrate from the root zone (1 m below surface) was modeled utilizing the Danish developed leaching model (NLES4) (Kristensen et al., 2008; Kronvang et al., 2008). The model was applied to a large number of combinations of soil types, crop types, climate, etc., and the N-leaching results were extrapolated to field block level within the catchment of the Horsens estuary based on field block information in agro-statistical data and climatic data for the agro-hydrological year of 2005 (1/4/2005 to 31/3/2006). For the agro-hydrological years 2000 to 2004, distributed data for nitrate leaching was estimated using agro-statistical data for 2005 because no specific regional data was available for 2000-2004. However, specific climate data for the years 2000 to 2004 was applied in the estimation of nitrate leaching. Nitrogen

Comment [dimul24]: Can drain runoff completely be regarded as runoff from groundwater, or will it be necessary both hydrologic ally and with respect to nutrient loads to accept tile and ditch drains as representing a mixture “zone” that include both surface and subsurface runoff components? If so, how much

Comment [d25]: Please specify a reference to "field block data in agro-statistical data" retention in groundwater was estimated by the differences between modeled net outlet of total N to surface waters from diffuse sources and the nitrate leaching from the root zone of the entire catchment.

For total P the situation is different as P concentrations are often up to one order of magnitude higher in deeper anaerobic aquifers compared to shallow aerobic aquifers and the phosphorus sources in anaerobic groundwater is generally natural. While the sources and transport of the different N-species are generally quite well known, the sources and transport of the various components of the measured total P are still poorly understood for subsurface as well as surface waters (Kronvang et al., 2007). As the major part of phosphorus in groundwater is natural it is neither relevant nor possible to derive a groundwater threshold value to control the anthropogenic input.

4 Results

4.1 Measured and modelled data from surface and subsurface waters

Nitrogen and phosphorus monitoring data for subsurface waters (suction cups, tile drains and monitoring wells) and surface waters (streams and estuary) are shown for comparison in Table 1 and 2, respectively. Model simulated concentrations for total N are compared to measured concentrations in Table 2 for the two gauged sub-catchments and for the Hansted stream in Fig. 3. The simulated concentrations fort the Bygholm stream is not as good as for the Hansted stream, but is still quite good (Nash-Sutcliff = 0.49), and as the model has not been calibrated on the measurements we consider it as a validation of our model setup.

Comment [d26]: What are the threshold values form which Nash-Sutcliff values are defined as "quite good". Why has the model not been calibrated before its application.

4.2 Development and current status for nutrient sources, loadings and sinks

The average land based nitrogen load to the estuary was 1770 metric tons yr^{-1} between 1984 and 1992 corresponding to an average weighted concentration in the streams of 11.1 mg N l^{-1} (Fig. 4). This concentration is 8 – 10 times higher than the estimated natural background loss. From 1993 the effects of abatement measures for nitrogen losses in agriculture become visible as nitrogen concentrations were decreasing in the freshwater discharge to the estuary reaching 5.1 mg N l^{-1} in 2009 (the simulated annual average for the investigated baseline period 2000-2005 is 6.2 mg N l^{-1}), Fig. 4. This concentration includes nitrogen from diffuse sources as well as point sources (sewage).

Comment [d27]: Provide a reference why one can assume that diffuse contributions from land via the shores are neglectable.

The most important source of N was agriculture, being responsible for 65% of the total N loading (Table 3). The average N loss from agricultural areas in the catchment amounted to 56 kg ha^{-1} yr^{-1} during the period 2001 to 2005, the period with the most detailed data and modeling. The second most important N source is the estimated loss of N from natural background sources, which amounts to 17%. The loadings from point sources in the catchment and marine fish farming amounted to 105 metric tons N, or only 9.7% of the total N loading (Table 3). Atmospheric deposition of N directly on the estuarine waters amounts to 8.7% of the total N loading.

Total phosphorous loadings to the Horsens estuary were, on average 95 metric tons P yr^{-1} from 1984 to 1987 (Fig. 4). Introduction of tertiary treatment of wastewater caused a sharp decline in 1988 and loadings continued to decline until 1995, reaching an average loading of 28 metric tons P yr^{-1} during 1995 to 2006 (Fig. 4). The average total P loading to the Horsens estuary amounted to 23.4 metric tons P during the period 2001 to 2005. The diffuse sources of P (background, agriculture and scattered dwellings) were the dominant source amounting to 16.2 metric tons P, or 69% of the total loading (Table 3). The second most important P source was urban runoff (15%), discharges from waste water treatment plants (8%), and fish farming in the estuary (6%).

The modeled average annual N-leaching from the root zone (1 m depth) on agricultural land in three sub-catchments to the Horsens Estuary is shown in Table 5. The N-leaching varies from year to year and from sub-catchment to sub-catchment, being dependent on factors such as climate, soil types, crop types and the application of chemical fertilizer and manure. The total annual N-leaching from both agricultural and non-agricultural land in the entire catchment to the Horsens Estuary is shown in Table 6. The N-leaching varies considerably from year to year being lowest in 2005 (1390 metric tons N) and highest in 2001 (3384 metric tons N). The N-transport in the streams was considerably lower than the modeled N-leaching (Table 6) due to N-removal in groundwater within the catchment. The average annual N-removal in groundwater amounts to 53% of the average annual N leached from the root-zone within the catchment with only slight variations from year to year (Table 6). The modeled N-removal in surface waters (streams, lakes, and wetlands) within the catchment is much lower than the N-removal in groundwater (Table 6). The average annual N-removal in surface water is 21% of the gross emission from diffuse and point sources for the study period. The resulting modeled annual N-loading and flow-weighted concentrations in inlet waters from diffuse sources to the Horsens Estuary are shown in Table 6. These flow-weighted concentrations vary between 4.4 and 6.0 mg N L^{-1} in the period 2000 to 2005 (the period with detailed modeling). The average annual N-fluxes from fields to the estuary are shown in Fig. 5. An average 64% of the N-emission from the diffuse sources is removed during the transport from field to estuary.

4.3 Relationships between nutrient loads and environmental status of Horsens estuary

Fig. 7 illustrate the relation between observed and modeled DIN concentrations in the estuary and show that 70% of the variability in DIN-concentrations can be explaned by N-loadings and wind stress. The nutrient concentrations in the estuary have declined concurrent with the decrease in loadings (Fig. 4 and 6). Decreasing chlorophyll concentrations were also observed in the inner part of the estuary for the spring period (March to June) from 1985 to 1992 following the drop in phosphorous loadings (data not shown). This is in agreement with indications of phosphorous as the primary limiting nutrient in the spring. However, in the outer part of the estuary and for the late summer period (July to October) the chlorophyll concentrations did not respond to the decrease in loadings and nutrient concentration. Water clarity improved from 1985 to 1995 in both parts of the estuary in the spring period (April to June). The diffuse attenuation coefficient (K_d) decreased from 1.15 m^{-1} to 0.55 m^{-1} in the inner part of estuary and from 0.81 to 0.33 m^{-1} in the outer part. Again, this is most likely a response to the lower phosphorous loadings and a general pattern observed in Danish estuaries where conditions in the spring are more directly influenced by loadings compared to conditions later in the summer where avialable nutrients are more governed by internal processes, e.g. release from the sediments. Since 1995 K_d has shown an increasing trend for the spring period and K_d-values from July to September have been variable with average values of 0.78 and 0.50 m^{-1} in the inner and outer part, respectively (Table 4), but no trends have been observed. Similarly, no positive developments have been observed for underwater vegetation (mainly eelgrass, *Zostera marina*, L.), which reached the lowest levels during the period 2000 to 2003. However, some improvements have been seen in 2007 to 2008 (Markager et al., 2010). Thus, despite significant reductions in nutrient loads and concentrations we only observe minor positive effects on the biological components in the

Comment [d28]: Have you checked for normal distribution?
What were the reasons (apart from statistical purposes) to regard the values of the years 1993 and 1994 as outliers, i.e. what physical / biological conditons may have different during these years?
Why were the years 2003 to 2006 save for validation. Visually it seems that the the years represents a period where the model overestimates.

Comment [d29]: Please explain or provide a reference that explains how wind stress regulates DIN-concentrations.

Comment [d30]: I suggest to provide a reference here ecosystem. Major improvements would require that the former eelgrass meadows where back and that water clarity and oxygen conditions have improved substantially (see 5.2 for a discussion of good ecological status).

Several mechanisms can explain the lack in biological response to the decrease in loads. A pool of nutrients in the sediment is probably the reason for a delay in the decline in nutrient concentrations. Generally positive residuals for nitrogen, i.e. observed concentrations that are higher than expected from the models, are seen over nine years from 1992, when nitrogen concentrations in the streams begin to drop, and until 2001 (Fig. 6). This could indicate a transition period where a positive net nitrogen flux out of the sediments is important. Another important mechanism is resuspension of sediment particles after the former underwater meadows of eelgrass are lost. A third factor can be derived from Fig. 8 showing the relationship between concentrations of inorganic nutrients and number of days with nutrient limitations, where nutrient limitations is assumed to occur at 14 µg DIN L^{-1} and 6.2 µg DIP L^{-1}. These values are equivalent to K_m-values for growth in a Michaelis-Menten expression of 1 µmol L^{-1} for DIN and 0.2 µmol L^{-1} for DIP based on values given by MacIsaac and Dugdale (1969), Eppley et al. (1969), Falkowski (1975) and Quile et al. (2011). For average DIN-concentrations (May-October) above 35 µg L^{-1} the percent of the time with N-limitation is rather constant (Fig. 8). Thus, DIN is in surplus and does only occasionally limit the growth of phytoplankton during the growth season, particularly in the inner part of the estuary (Fig. 8a). Only when the average DIN concentrations fall below about 35 µg L^{-1} will N-limitation become significant. This pattern indicates that the reductions in N-loads have removed a surplus of nitrogen in the estuary, but have until recently not been sufficient to introduce significant N limitation of phytoplankton growth. A similar figure for P shows a more linear

Comment [d31]: Please provide some corresponding evidence from times series data or references on nutrient concentrations in sediments

Formatted: English (U.K.)

Comment [d32]: I strongly suggest to start with discussion of these references to ensure that the reader understands that the authors have based there definition / derivation of the values "14 µg DIN L^{-1} and 6.2 µg DIP L^{-1}" have been derived from these references.

Formatted: English (U.K.)

increase in the time period with increasing limitation when average concentrations decline (Fig. 8b), and the inner and outer part of the estuary have approximately the same concentrations of DIP (Table 4).

4.4 Maximum acceptable N and P loads

Maximum acceptable total loads were defined on the basis of Fig. 8 and the assumption that nutrient limitation of phytoplankton growth is necessary during most of the growth season in order to achieve good ecological status (see 5.2 for a discussion of good ecological status). We find it necessary to apply a "dual-nutrient reduction strategy" wherein both N and P loads are reduced (Boynton and Kemp, 2008; Conley et al., 2009) in order to ensure good ecological status, and we have defined the average DIN and DIP concentrations where nutrient limitations occur during 2/3 of the growth season as a reasonable threshold (Fig. 8). The corresponding threshold values are 21 µg DIN L^{-1} and 7 µg DIP L^{-1} for the inner and outer estuary, respectively (Fig. 8a and 7b). Once we have defined the target value, the corresponding loads can be calculated from the empirical models assuming that climatic variable in models are equal to their long term mean values. These are a N-load of 560 metric tons yr^{-1} and a P-load of 13 metric tons P yr^{-1}. These loadings result in estimated DIN concentrations of 20 and 5.3 µg N L^{-1} for the inner and outer part of the estuary, respectively (Table 4). Thus, N-limitation will occur during 2/3 of the time (May to October) in the inner part and for about 95% of the time in the outer part. The estimated DIP concentrations corresponding to a total P load of 13 tons yr^{-1} to the estuary are 6.0 and 6.2 µg P L^{-1} for the inner and outer parts, respectively, which are close to the values resulting in nutrient limitation for 2/3 of the time from March to July. Please note that the concentrations for DIN (20 and 5.3 µg N L^{-1}) and DIP (6.0 and 6.2 µg P L^{-1}) are mean values over the season. Thus,

Comment [d33]: Please provide an explanation, how these specific threshold values have been set / defined higher concentration, allowing nutrient-replete growth of phytoplankton, will still occur for approximately 1/3 of the time.

The considerations above do only take DIN and DIP into account despite the fact that dissolved organic matter is by far the largest pool of nutrients, e.g. is the ratio of TN:DIN about 150 (Fig. 4 and 6a). However, dissolved organic nitrogen (DON) is not readily taken up by phytoplankton, and is mainly used indirectly after mineralization of DON by bacteria. The concentrations of both inorganic and organic N and P are determinated by loadings, biological processes and mixing with the marine end member. On an annual scale the estuary is a reactor transforming DIN (approximately 80% of the loadings) to DON (Stedmon et al. 2006; Markager et al. 2011).

An alternative method to define the target values for good ecological status is to use the empirical models to calculate concentrations for TN and TP with the values for background loadings. These will theoretically give the TN and TP concentrations at pristine conditions. However, the empirical models are then used for scenarios with loads far outside the range used for setting up the models and the outcome is therefore uncertain. For TN the estimated pristine concentration is 398 µg L^{-1}, when using the politically defined practice of accepting a 26% deviation from pristine conditions (Table 4). The corresponding load would be 743 tons N yr^{-1}, or 33% higher than the above mentioned 560 tons N yr^{-1}, however, given the uncertainty the two values are in reasonable agreement. For TP the model show a low sensitivity between loadings and concentration and estimated pristine concentrations are so high than an addition of 26% will bring them above the present concentrations, which clearly do not support a good ecological status. Thus, this approach does not work for TP. The reason for the low sensitive of the empirical model with respect to TP is probably a high amount of stored phosphorus in the sediments.

4.5 Calculated groundwater and stream threshold values and groundwater chemical status in the catchment of Horsens Estuary

The estimated maximum acceptable N and P loads (560 and 13 ton) required to ensure a good ecological status of the Horsens Estuary, were estimated in the previous section. These loads correspond to 52 and 56% of the annual average total N and P loads to the estuary for the period 2000 to 2005, respectively. To meet these reduction targets we calculate the following threshold values in the two possible scenarios described previously.

4.5.1 Reduction targets and threshold values – scenario 1.

The first scenario assumes that all reduction targets for N and P is directed against the diffuse sources in the catchment (Table 7). The resulting total N and P concentration in inlet freshwater to the estuary are calculated to 2.9 and 0.084 mg L^{-1}, respectively. The corresponding groundwater threshold value for total N in aerobic groundwater in the catchment is calculated to 6.0 mg L^{-1}. No groundwater threshold value in the catchment can be calculated for P as other diffuse sources such as soil erosion and stream bank erosion are important transport pathways, which currently are not completely quantified.

4.5.2 Reduction targets and threshold values - scenario 2.

In the second scenario we are imposing reduction targets on point sources, direct atmospheric deposition (emission from agriculture of ammonia), and diffuse sources (Table 8). The resulting inlet concentration in inlet freshwater to the estuary is 3.1 mg L^{-1} and the corresponding groundwater threshold value of N is calculated to 9.3 mg L^{-1} – thus being considerably higher than in scenario 1. The reason is that reduction in point sources, direct loads, and targeted mitigation measures such as restored wetlands and uncultivated buffer zones will assist in reducing the loadings to the estuary. The scenario 2 calculations for P show that the reduction target for the estuary can be achieved in a longer term perspective by introducing targeted mitigation measures.

The calculated stream and groundwater threshold values for the two scenarios are compared to current total N and P concentrations in Table 9. Note that the nitrate-N concentrations in streams is about 89 % of the total N concentration based on measurements at monitoring stations, hence the threshold value (TV) for nitrate-N is also 89 % of the TV for total N given here. The TV for nitrate-N in groundwater equals in practice the TV for total-N based on measurements in monitoring wells. The modeled groundwater concentrations are recharge-weighted. The mean concentration of a sufficient number of monitoring wells in aerobic groundwater should equal this number if aerobic groundwater represents the same recharge period as the modeled baseline period i.e. 2000 to 2005.

5 Discussion

5.1 Estimate of total N and P loads from gauged and ungauged catchments to the Horsens Estuary

The model simulations of nitrogen leaching and the modeled gross and net nutrient emissions are believed to be of relatively high precision as the models applied are empirical models developed based on the national monitoring data from agricultural fields in agricultural

Comment [d34]: What is relatively high precision? Moreover one could argue that models developed on national data sets are less well adopted to regional. Of course from the applicability point of view it would be a big advantage, if a model developed on national data sets perform well in catchments of various size and location.

catchments (Grant et al., 2007) and stream monitoring data from 80 catchments (Windolf et al., 2011). This conclusion is corroborated by the good fit to the measured stream concentrations in the gauged sub-catchments (Table 2). The simulated nitrogen concentration in the Hansted catchment equals the measured values, whereas the simulated values in the Bygholm catchments are slightly lower than the measured values. The latter deviation is the cause of the slightly lower estimate for the annual N loading to the estuary based on simulated values for the Bygholm and Hansted catchments (1001 metric tons for the period 2001 to 2005) as compared to the estimate using measured N loadings for these two catchments (1086 metric tons). Of course this will also affect the final computed threshold values for total N (Table 9). A previous inter-comparison of model estimates have shown that the precision of N modeling in catchments is rather high, whereas P modeling estimates currently have a poor precision (Kronvang et al., 2009a).

Comment [d35]: This sounds like a better argument

Comment [d36]: Do the authors or the authors of the 2009 paper provide any qualified estimations, why this might be the case?

Comment [d37]: Check against reference list. Should *a* be added as a suffix?

5.2 Estimate of maximum acceptable loads

Comment [d38]: Good chapter, but I am still missing the identification of the ecologically defined target values from the theoretical point of view. The values specified in figure 8

A key issue for management of an estuary is to establish maximum acceptable loads. An assessment of this involves the definition of target values for one or several parameters in the estuary that describe good ecological status. Then, models for quantitative relationships between loads and these parameters are needed to estimate the maximum acceptable loads required to reach these target values.

Recent research has demonstrated that dual-nutrient (N, P) reduction strategies are needed to alleviate eutrophication in estuaries and other coastal waters in the land-sea continuum (Boynton and Kemp, 2008; Conley et al., 2009; Paerl et al., 2009), and that the Redfield ratio for N and P in marine waters (16:1, molar) cannot be considered a universal optimal ratio between N and P, but rather an average of species-specific N:P ratios (Klausmeier et al., 2004; Ptacnik et al. 2010).

Comment [d39]: Check, as there is no reference of Paerl with more than 2 authors

Our approach has been to define good ecological status as average concentrations of inorganic nutrients, which ensure nutrient limited phytoplankton growth in 2/3 of the growth season, taking into account the natural seasonal cycle where phosphorous is limiting in the spring and nitrogen is limiting later in the growth season.

The choice of 2/3 of the growth season may be debatable. Moreover, it is known that the K_m-value for growth of phytoplankton varies between species (e.g. Falkowski, 1975) and that growth rates are more closely coupled to the internal cell concentrations than to external concentrations. However, we still find that the selected approach is based on reasonable ecological rationales and that it gives a good indication of the nutrient concentration levels that ensure an acceptable ecological status of the estuary. As recognized by Duarte et al. (2009) the definition of target loads and concentrations for achieving good ecological status of estuaries is probably the most challenging part of the restoration process. In the end the definition of good ecological status will always have a political dimension and our scientifically based definitions of good ecological status and implied targets for loadings can only be guidelines for the political decision process.

The use of empirical models for relationships between loads and nutrient concentrations in the estuary works well for nutrient concentrations. However, it is important to remember that empirical models describe the present conditions in the estuary and only have a time lag between loads and effects in the estuary of approximately one year. Thus, effects with a longer time lag and possible regime shifts (Scheffer, 2001) are not accounted for. This is presumably the reason why changes in water clarity and depth limits of eelgrass gives very weak models with low sensitivity (data not shown). This is most likely due to pools of nutrients stored in the sediments, which only slowly (presumably over decades) are released and emptied during a phase with decreasing loadings. Predicting these time lags and regime shifts, e.g. from the present phytoplankton dominated system back to an eelgrass dominated system, is extremely difficult but clearly a major scientific challenge for the coming years.

In conclusion, the empirical models applied here provide a reasonably good prediction of nutrient concentrations during changes in loadings within the range of loadings for which they are developed. Effects of changes in loadings significantly outside this range or for other regimes of the ecosystem are very uncertain. The lowest loadings in the data set encompass the predicted targets for N and P so the model are not used outside the data range. However, additional effects of processes with time lag of decades are not accounted for.

5.3 Management options for N and P in Horsens Estuary

The reduction targets for N (526 tons) and P (10.4 tons) can be accomplished by different mitigation measures in the catchment and introducing improved treatment of sewage water at point sources discharging either to freshwater or directly to the estuary. As described previously, we have developed two possible management options that could be introduced to reduce the N and P loadings to levels allowing good ecological status in the Horsens estuary.

The first scenario assumes that the entire N reduction is obtained by introducing mitigation measures, which reduce the N leaching from the root zone of agricultural fields. The inlet total N and total P concentration in freshwater discharging to the Horsens estuary has to be reduced from 6.2 to 2.9 mg L^{-1} and 0.15 mg L^{-1} to 0.084 mg L^{-1}, respectively, for obtaining good ecological status. The resulting model calculated threshold value of total N in the root zone and aerobic groundwater at and below a depth of one meter is 6.0 mg L^{-1}(equivalent to 26.5 mg L^{-1} NO_3^-) as an average for the entire catchment area (Table 7). However, the threshold value for total N under agricultural fields can be allowed to be higher (7.4 mg L^{-1} equivalent to 32.7 mg L^{-1} NO_3^-) because approx. one third of the catchment area is in a non-agricultural land cover category, with a low background concentration of total N in groundwater (< 1 mg L-1 in some areas, (Postma et al., 1991)) and streams (approx. 1.2 mg L^{-1}) (Kronvang et al., 2005). As phosphorus is derived via many hydrological pathways (leaching, erosion, and surface runoff) to surface waters (Kronvang et al., 2007) it is not possible to calculate a groundwater P threshold value with our current knowledge.

Our second reduction scenario for N and P involves reduction in discharges of nutrients from point sources, enhancing N and P retention processes in surface waters (reestablishing riparian wetlands, introducing buffer strips, etc.) and reductions in diffuse sources (Hejzlar et al., 2006; Hoffmann et al., 2009; Kronvang et al., 2009; Hoffmann et al., 2011). Such a catchment management plan allows the groundwater threshold value to be higher (average for entire catchment area: 9.3 mg N L^{-1}) than in the first scenario. The threshold N concentration under agricultural fields in the catchment is then calculated to 11.8 mg N L^{-1} (52 mg L^{-1} as nitrate). Note that the latter is above the U.S. as well as the European drinking water standards of 10 mg L^{-1} nitrate-N (~ 44 mg L^{-1} nitrate) and 50 mg L^{-1} nitrate, respectively. In such a case the drinking water standard will have to be applied as a threshold value according to European directives and guidelines. The second scenario for P seems to be enough to reduce the P-loadings to the required target and reach the corresponding threshold value of 0.084 mg L^{-1} for phosphorus in streams. This will, however, take some time, as some of the surface water management methods need a long period to work efficiently in reducing P (buffer strips, Table 7).

An additional management option for reduction of nutrient loadings to the estuary is linked to a spatial analysis of nitrogen sources within the catchment to Horsens estuary, where the catchment is divided into sub-catchments (Windolf et al., 2011). Lumped results of model calculations of gross N emissions and sinks within 27 sub-catchments are available for the Horsens Estuary catchment. Eight of these sub-catchments are located downstream of the larger lakes in the catchment (downstream from the two river monitoring stations) so management of N within agricultural production in this area will be most cost-effective as no natural N reduction takes place in lakes in these sub-catchments (Thodsen et al., 2009).

The management option chosen is to transform land use from agricultural land to forest land in this 154 km^2 sub-catchment. This will lead to a reduction of the N-loading to the estuary of 200 metric tons N per year. The remaining 326 metric tons N has to be removed from the catchment upstream the two larger lakes. An annual N retention of 13% of the incoming N load to the two lakes (Bygholm and Nørrestrand) has been calculated using the N retention model from Windolf et al. (2011). Thus, the N loading to these two lakes has to be reduced to 409 metric tons N per year. As the retention of N in groundwater and surface waters within the catchment upstream the two lakes amounts to around 60% of the N leached from the root zone, we can calculate that the threshold N concentration in upper groundwater can be allowed to be approximately 10% higher than the threshold value of 7.4 mg N/L under agricultural areas calculated in scenario 1.

5.4 Estimation of groundwater threshold values from maximum acceptable loads and different management options

It has been demonstrated through the previous sections that groundwater threshold values derived based on maximum acceptable loads to an associated aquatic ecosystem depend on technically and politically realistic management options to reduce nutrient loads to the ecosystem. Consequently, groundwater threshold values for nutrients derived to protect ecosystems will never be universal as drinking water standards often are. Ecological driven groundwater threshold values should always be derived for a specific geological, climatological and agricultural setting. Values derived for similar settings may, however, be used if data in given water bodies and ecosystems are insufficient for derivation of

Comment [d40]: To my mind the authors are mixing the “ecological driven groundwater thresholds” with “what achievable” for a given setting. In my opinion ecological driven thresholds should purely be defined according to the ecological needs of the target ecosystem. According to Acosta-Alba (Sustainability 2011, 3, 424-442; doi:10.3390/su3020424) **are** indicators environmental statistics that measure or reflect environmental status or change in condition. Thresholds are critical levels for these indicators; a threshold level representing the level beyond which a system undergoes significant change. If groundwater threshold values. Groundwater threshold values derived for a comparable setting should probably often be preferred to drinking water standards, which for example are currently used as the threshold value for nitrate by most European countries. The calculated groundwater thresholds in this paper are average annual flow (recharge)- weighted concentrations acceptable in aerobic groundwater discharging to streams in the catchment. As the water table and the upper aerobic groundwater zone are very shallow in the investigated catchment (< 5m), the aerobic groundwater generally recharged the aquifers within the last few years. Hence, average concentrations in a representative number of monitoring screens in the aerobic zone (if present) should not exceed the flow-weighted groundwater threshold values obtained by the conducted model simulations. Unfortunately, the number of monitoring wells in aerobic groundwater in the catchment is very small and several of them are probably screened across the redox boundary. The average total N concentration in aerobic groundwater calculated from monitoring wells in the aerobic zone is therefore not considered to be representative for aerobic groundwater in the catchment.

5.5 Groundwater chemical status

If a groundwater threshold value derived for protection of an associated ecosystem is ~~b~~reached in a given groundwater body, the groundwater body or part of a groundwater body has to be classified as having poor status. In the case of nitrogen for example it is necessary to evaluate the concentrations of the different nitrogen species separately for the aerobic and anaerobic parts of the groundwater bodies. This is important as nitrate, which represents practically the entire total N in aerobic groundwater, is reduced to the inactive harmless N_2 in anaerobic groundwater (e.g. Appelo and Postma, 2005). Consequently total N concentrations are typically more than an order of magnitude lower in the anaerobic zone than in the aerobic zone, and the anaerobic zone thus contributes relatively little to N loads.

Consequently, the general groundwater chemical status in the catchment based on nitrogen species should generally be assessed for the aerobic groundwater separately. Conceptual models of the extension of the aerobic groundwater and their role for surface water nitrogen loads as represented here (Fig. 2) should support the risk analysis. If data on aerobic groundwater are missing or scarce, measured stream nitrate or total N concentrations are useful indicators of the status of the shallow aerobic groundwater in the catchment, when wastewater and other nitrogen sources are taken into account. This is clearly illustrated when comparing results from Fig. 9 and Table 9. Fig. 9 leaves the impression that relatively few groundwater bodies have problems with nitrate, while data in Table 9 clearly demonstrate that nitrate concentrations are generally too high in the catchment. Hence, the conducted model simulations show that the groundwater chemical status based on nitrate concentrations in aerobic groundwater is poor below farm lands in general in the area, and that the quality of shallow aerobic groundwater in the catchment does not comply with European legislation.

Comment [d41]: Unless one recalls figure 1, the gauging stations are difficult to find

5.6 Implications for integrated monitoring and modeling of water bodies

The previous section clearly demonstrates that groundwater and surface water monitoring should be integrated in order to obtain as much information as possible on the chemical status of both water body types, and in order to derive meaningful groundwater threshold values for protection of associated and dependent aquatic and terrestrial ecosystems. As the ecological status of surface waters depends on the nutrient loadings and the seasonality in nutrient loadings, water quality monitoring programs should provide the necessary data to calculate and simulate these by coupled groundwater and surface water models, not least when possible climate change impacts have to be assessed (Andersen et al., 2006; Sonnenborg et al., 2011). In addition, reliable models and design of efficient monitoring programs for assessment of groundwater impacts on ecosystems require a sound understanding of the site specific hydrogeological, physical, and chemical conditions controlling the groundwater – surface water interaction (Dahl et al., 2007, Dahl and Hinsby, in press). This challenges the traditional and still very relevant groundwater monitoring of major aquifers, which is targeted drinking water interests. Furthermore, it may also challenge surface water monitoring traditions, as models being able to simulate runoff and nutrient concentrations with a high spatial and temporal variation and coverage are needed, and they require reliable monitoring data for calibration.

5.7 Climate change impact on N and P loadings to coastal ecosystems

Before concluding this work a short note on the possible effect of projected climate change on groundwater threshold values in the investigated study area is in place. Much research is currently undertaken in order to assess the projected climate change impact on e.g. the hydrological cycle, globally. Previous work has indicated that winter precipitation and hence nutrient loadings to coastal waters may increase in Denmark (Andersen et al., 2006; Jeppesen et al., 2009a,b; van Roosmalen et al., 2009; Aquarius, 2011; Sonnenborg et al., 2011) although significant uncertainties exist e.g. due to changes in crops and farming practices (Olesen et al., 2007). Furthermore, while increased temperatures are expected to increase crop yields in the North Sea and Baltic Sea regions (Aquarius, 2011), the increased temperatures will render coastal ecosystems more prone to harmful algal blooms (Paerl and Huisman, 2009) and hypoxia as mineralizations accelerates with higher temperatures. In such a scenario groundwater threshold values will have to be lower than the values derived in this paper. Hence, for Denmark and the other countries in the region the measures, which are

Comment [d42]: Check against reference list implemented to assure good chemical and ecological status of water bodies, may not be sufficient in the future as projected climate change may work against these. The present paper sets the scene and establishes the needed knowledge base for integrated understanding of the Horsens estuary and catchment system for the assessment of climate change impacts on groundwater threshold values and chemical status. This issue is the subject of a companion paper (in prep.).

6 Conclusion

As a result of the intensive agriculture in Denmark the majority of Danish coastal waters have poor ecological status. Hence, the development of catchment or river basin management plans for reduction of nutrient loads and determination of threshold values in groundwater, streams, and estuaries are becoming increasingly important. The present study analyses and presents 1) the historical and current nutrient loadings for the investigated Horsens Estuary 2) the current ecological conditions of the estuary and 3) necessary reductions in nutrient loadings for obtaining a good ecological status in the estuary applying a suite of empirical loading-response models. We estimate that the total N and P annual loads for the investigated baseline period (2000 to 2005) should be reduced to 560 and 13 ton, respectively, corresponding to 52 and 56% of the annual average for the investigated baseline period. Using different scenarios we demonstrate that, especially the groundwater threshold values or maximum acceptable concentrations are quite sensitive to the choice of mitigation measures and management options in the catchment. Depending on the selected management scenario we estimate that groundwater threshold values for total N vary between 6.0 and 9.3 mg/L, while the corresponding stream threshold values vary between 2.9 and 3.1 mg/L. As the current modeled average concentrations in shallow aerobic groundwater and streams are 15 and 6.2 mg L^{-1}, respectively, our investigation clearly shows that groundwater and stream threshold values are breached in the catchment. Hence, the major part of the shallow aerobic groundwater in the catchment to Horsens Estuary is of poor chemical status due to farming practices and does not comply with the European Water Framework and Groundwater Directives. To obtain good chemical status for shallow aerobic groundwater in the investigated catchment, our data show that the average total N concentrations should be lowered to approximately half (40 to 62% - depending on the applied management option) of the present concentration. These reductions correspond to NO_3-N threshold values in the range of 6 – 9 mg L^{-1} (or 27 to 41 mg L^{-1} of nitrate) assuming that the nitrate species constitute the entire total N in shallow aerobic groundwater. According to our evaluation, the flow-weighted annual average concentration of total P in streams in the catchment should be lowered from the present 0.15 to 0.084 mg/L. However, the present study indicates that it is not relevant to establish groundwater threshold values for total P in the investigated catchment as the elevated concentrations apparently occur only in anaerobic groundwater due to dissolution from natural sources, and a major and unknown part of the total P in streams originates from brink erosion. The transport of total P is, however, not as well understood as the transport of total N and should be investigated further. It is interesting to note that one of the presented management scenarios would allow aerobic groundwater nitrate concentrations below farm lands even above drinking water standards if focusing solely on the good status objective for the estuary. However, such high concentrations would jeopardize the chemical status of groundwater used for drinking water, and the ecological status of ecosystems in the catchment such as lakes and wetlands. Hence, an integrated assessment of acceptable loads and thresholds for both coastal waters and surface and subsurface waters in the catchment is imperative, when thresholds have to cover other relevant ecosystems in a catchment such as lakes and protected terrestrial ecosystems. The threshold values derived in this study to ensure good ecological status of the Horsens estuary may not ensure good ecological status for all ecosystems in the catchment. Furthermore, climate change impacts will most probably require lower groundwater and stream threshold values in the future to ensure good ecological status of associated aquatic ecosystems.

Acknowledgment

This article is an outcome of the EU Interreg IVB project CLIWAT. It has been co-funded by the North Sea Region Program 2007-2013 under the ERDF of the European Union, and supported by the Danish Strategic Research Foundation projects MONITECH (104-08-0050) and MAFIA (09-067259). We acknowledge Lauren Paige Seaby for linguistic assistance.

References

Andersen, H. E., Kronvang, B. and Larsen, S. E.: Development, validation and application of Danish empirical phosphorus models, J. Hydrol., 304, 355-365, 2005.

Andersen, H. E., Kronvang, B., Larsen, S. E., Hoffmann, C. C., Jensen, T. S. and Rasmussen, E. K.: Climate-change impacts on hydrology and nutrients in a Danish lowland river basin, Sci. Total Environ., 365, 223-237, 2006.

Appelo, C. A. J. and and Postma, D.: Geochemistry, groundwater and pollution, Balkema, Leiden, 2005.

Aquarius: Main findings from Aquarius – Farmers as water managers in the North Sea region, Report from the Aquarius project, the Interreg IVB North Sea Region Programme, Knowledge Centre for Agriculture, Aarhus, 2011.

BLST: Vandplan - Hovedopland 1.9 Horsens, Danish Agency for Spatial and Environmental Planning, Ministry of the Environment, Copenhagen, 1-206 pp., 2010.

Boynton W.R. and Kemp, W. M.: Estuaries, in: Nitrogen in the Marine Environment, Capone, D. G. e. al. (Eds.), Elsevier, Burlington, 809-866, 2008.

Boynton, W. R., Garber, J. H., Summers, R. and Kemp, W. M.: Inputs, Transformations, and Transport of Nitrogen and Phosphorus in Chesapeake Bay and Selected Tributaries, Estuaries, 18, 285-314, 1995.

Broadhurst, D., Goodacre, R., Jones, A., Rowland, J.J. and Kell, D.B.: Genetic algorithms as a method for variable selection in multiple linear regression and partial least squares regression, with applications to pyrolysis mass spectrometry. –Ana. Chim. Acta, 348 (1-3), 71-86, 199~~7~~7.

Camargo, J. A. and Alvaro, A.: Ecological and toxicological effects of inorganic nitrogen pollution in aquatic ecosystems: A global assessment, Environ. Int., 32, 831-849, 2006.

Cloern, J.: Our evolving conceptual model of the coastal eutrophication problem, Mar. Eco. Prog. Ser., 210, 223-253, 2001.

Conley, D. J., Kaas, H., Mohlenberg, F., Rasmussen, B. and Windolf, J.: Characteristics of Danish estuaries, Estuaries, 23, 820-837, 2000.

Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., Lancelot, C. and and Likens , G. E.: Controlling Eutrophication: Nitrogen and Phosporus, Science, 323, 114-115, 2009.

Comment [d43]: Reference not used

Comment [d44]: Check against use in text (Broadhurst 1997)

Comment [d45]: Reference not applied

Comment [d46]: Reference not used

Conley, D. J., Markager, S., Andersen, J., Ellermann, T. and Svendsen, L. M.: Coastal eutrophication and the Danish National Aquatic Monitoring and Assessment Program, Estuaries, 25, 848-861, 2002.

Dahl, M. and and Hinsby, K.: Typology of groundwater – surface water interaction (GSI) typology, in: Groundwater and Ecosystems, Selected papers on Hydrogeology 17, Ribeiro, L. and et al. (Eds.), Taylor and Francis, Oxon, in press.

Dahl, M., Nilsson, B., Langhoff, J. H. and and Refsgaard, J. C.: Review of classification systems and new multi-scale typology of groundwater-surface water interaction, J. Hydrol., 344, 1-16, 2007.

Diaz, R. J. and Rosenberg, R.: Spreading dead zones and consequences for marine ecosystems, Science, 321, 926-929, 2008.

Duarte, C. M., Conley, D. J., Carstensen, J. and and Sánchez-Camacho, M.: Return to Neverland: Shifting baselines affect eutrophication restoration targets, Estuar. Coast., 32, 29-36, 2009.

Eppley, R. W., Rogers, J. N. and and McCarthy, J. J.: Half-saturation constants for uptake of nitrate and ammonium by marine phytoplankton, Limnol. Oceanogr., 14, 912-920, 1969.

EU, directive 2000/70/EC ("The EU Water Framework Directive"), 2000.

EU, directive 2006/118/EC ("The EU Groundwater Directive"), 2006.

EU, directive 2009/90/EC, ("The Monitoring Directive"), 2009

Falkowski, P. G.: Nitrate uptake in marine phytoplankton: Comparison of half saturation constants from seven species., Limnol. Oceanogr., 20, 412-417, 1975.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R. and Vorosmarty, C. J.: Nitrogen cycles: past, present, and future, Biogeochemistry, 70, 153-226, 2004.

Gustafsson, K. E. and and Bendtsen, J.: Elucidating the dynamics and mixing agents of a shallow fjord through age tracer modeling. , Est Coast Shelf Sci, 74, 703-716, 2007.

Hagg, H. E., Humborg, C., Morth, C. M., Medina, M. R. and Wulff, F.: Scenario Analysis on Protein Consumption and Climate Change Effects on Riverine N Export to the Baltic Sea, Environ. Sci. Technol., 44, 2379-2385, 2010.

Comment [d47]: Reference not used in the text

Hansen, B., L., Thorling, L., Dalgaard, T. and and Erlandsen, M.: Trend Reversal of Nitrate in Danish Groundwater – a Reflection of Agricultural Practices and Nitrogen Surpluses since 1950, Env Sci Tech, 45, 228-234, 2011.

Hejzlar, J., Samalova, K., Boers, P., and Kronvang, K.: Modeling phosphorus retention in lakes and reservoirs. - Water, Air, Soil Pollut., 6, 487-494, 2006.

Hejzlar, J., Anthony, S., Arheimer, B., Behrendt, H., Bouraoui, F., Grizzetti, B., Groenendijk, P., Jeuken, M. H. J. L., Johnsson, H., Lo Porto, A., Kronvang, B., Panagopoulos, Y., Siderius, C., Silgram, M., Venohr, M. and Zaloudik, J.: Nitrogen and phosphorus retention in surface waters: an inter-comparison of predictions by catchment models of different complexity, J. Environ Monitor., 11, 584-593, 2009.

Comment [d48]: Reference not used

HELCOM: HELCOM Baltic Sea Action Plan, HEE, 101 pp., 2007.

Henriksen, H. J., Troldborg, L., Nyegaard, P., Sonnenborg, T. O., Refsgaard, J. C. and and Madsen, B.: Methodology for construction, calibration and validation of a national hydrological model for Denmark. , J. Hydrol., 280, 52-71, 2003.

Hinsby, K. and J¢rgensen, L. F.: Groundwater monitoring in Denmark and the Odense Pilot River Basin in relation to EU legislation, in: Groundwater Monitoring, Ph.Quevauviller, AM Fouillac, J.Grath and R.Ward (Eds.), Wiley, Hoboken, 207-224, 2009.

Hinsby, K., de Melo, M. T. C. and Dahl, M.: European case studies supporting the derivation of natural background levels and groundwater threshold values for the protection of dependent ecosystems and human health, Sci. Total Environ., 401, 1-20, 2008.

Hoffmann, C. C., Kjaergaard, C., Uusi-Kamppa, J., Hansen, H. C. B. and Kronvang, B.: Phosphorus Retention in Riparian Buffers: Review of Their Efficiency RID B-6714-2011, J. Environ. Qual., 38, 1942-1955, 2009.

Hoffmann, C. C., Kronvang, B. and Audet, J.: Evaluation of nutrient retention in four restored Danish riparian wetlands, Hydrobiologia, 674, 5-24, 2011.

Jeppesen, E., Kronvang, B., Meerhoff, M., Sondergaard, M., Hansen, K. M., Andersen, H. E., Lauridsen, T. L., Liboriussen, L., Beklioglu, M., Ozen, A. and Olesen, J. E.: Climate Change Effects on Runoff, Catchment Phosphorus Loading and Lake Ecological State, and Potential Adaptations, J. Environ. Qual., 38, 1930-1941, 2009.

Jeppesen, E., Kronvang, B., Olesen, J. E., Audet, J., Sondergaard, M., Hoffmann, C. C., Andersen, H. E., Lauridsen, T. L., Liboriussen, L., Larsen, S. E., Beklioglu, M., Meerhoff,

Comment [d49]: Reference not used

M., Ozen, A. and Ozkan, K.: Climate change effects on nitrogen loading from cultivated catchments in Europe: implications for nitrogen retention, ecological state of lakes and adaptation, Hydrobiologia, 663, 1-21, 2011.

Klausmeier, C., Litchman, E., Daufresne, T. and Levin, S.: Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton, Nature, 429, 171-174, 2004.

Kronvang, B., and Bruhn, J.A.: Choice of sampling strategy and estimation method when calculating nitrogen and phosphorus transport in small lowland streams, Hydrol Process, 10, 1483-1501, 1996.

Kronvang, B., Grant, R., Larsen, S. E., Svendsen, L. M. and Kristensen, P.: Non-Point-Source Nutrient Losses to the Aquatic Environment in Denmark - Impact of Agriculture, Marine and Freshwater Research, 46, 167-177, 1995.

Kronvang, B., Jeppesen, E., Conley, D. J., Sondergaard, M., Larsen, S. E., Ovesen, N. B. and Carstensen, J.: Nutrient pressures and ecological responses to nutrient loading reductions in Danish streams, lakes and coastal waters, J. Hydrol., 304, 274-288, 2005.

Kronvang, B., Vagstad, N., Behrendt, H., Bogestrand, J. and Larsen, S. E.: Phosphorus losses at the catchment scale within Europe: an overview, Soil Use Manage., 23, 104-116, 2007.

Kronvang, B., Andersen, H. E., Borgesen, C., Dalgaard, T., Larsen, S. E., Bogestrand, J. and Blicher-Mathiasen, G.: Effects of policy measures implemented in Denmark on nitrogen pollution of the aquatic environment, Environ. Sci. & Policy, 11, 144-152, 2008.

Comment [d50]: Check against use in text. No 2011 reference of Jeppessen used

Kronvang, B., Hoffmann, C.C., and Dröge, R.: Sediment deposition and net phosphorus retention in a hydraulically restored lowland river-floodplain in Denmark: combining field studies with laboratory experiments, Marine and Freshwater Research, 60, 638-646, 2009a.

Kronvang, B., Behrendt, H., Andersen, H. E., Arheimer, B., Barr, A., Borgvang, S. A., Bouraoui, F., Granlund, K., Grizzetti, B., Groenendijk, P., Schwaiger, E., Hejzlar, J., Hoffmann, L., Johnsson, H., Panagopoulos, Y., Lo Porto, A., Reisser, H., Schoumans, O., Anthony, S., Silgram, M., Venohr, M. and Larsen, S. E.: Ensemble modelling of nutrient loads and nutrient load partitioning in 17 European catchments, Journal of Environmental Monitoring, 11, 572-583, 2009b.

Kronvang, B., Hoffmann, C. C. and Droge, R.: Sediment deposition and net phosphorus retention in a hydraulically restored lowland river floodplain in Denmark: combining field and laboratory experiments RID B-6714-2011, Marine and Freshwater Research, 60, 638-646, 2009c.

Kristensen, K., Waagepetersen, J., Børgesen, C.D., Vinther, F.P., Grant, R. and Blicher-Mathiesen, G.: Reestimation and further development in the model N-LES N-LES3 to N-LES4. DJ F PLANT SCIENCE, 139, Faculty of Agricultural Sciences, Aarhus University, Denmark, 2008.

Kyllingsbæk, A.: Landbrugets husholdning med næringsstoffer 1900-2005 - Kvælstof Fosfor Kalium. DJF Markbrug, 18, 1-77, Faculty of Agricultural Sciences, Aarhus University, Denmark, 2008.

Comment [d51]: Check against use in text, as one or more of Kronvang et al. 200 has not been used and hence might be obsolete e in the reference list

Comment [d52]: Not used in text

Larsen, M. A. D., Soegaard, H. and Hinsby, K.: Temporal trends in N & P concentrations and loads in relation to anthropogenic effects and discharge in Odense River 1964-2002, Hydrology Research, 39, 41-54, 2008.

Comment [d53]: Not used in text

Larsen, S. E., Kronvang, B., Windolf, J. and Svendson, L. M.: Trends in diffuse nutrient concentrations and loading in Denmark: Statistical trend analysis of stream monitoring data, Water Science and Technology, 39, 197-205, 1999.

Comment [d54]: Not used in text

MacIsaac, J. J., and Dugdale, C.: The kinetics of nitrate and ammonia uptake by natural populations of marine phytoplankton, Deep-Sea Res, 16, 45-57, 1969.

Markager, S., Storm, L. M. and and Stedmon, C. A.: Limfjordens miljøtilstand 1985 til 2003. Sammenhæng mellem næringsstoftilførsler, klima og hydrografi belyst ved empiriske modeller. Report no. 577, , National Environmental Research Institute, Aarhus University, Aarhus, Denmark, 2006.

Comment [d55]: Not used in text

Markager, S., Bassompierre, M. and Petersen, D. J.: Analyse af Milj¢tilstanden i Horsens Fjord fra 1985 til 2006 - empirisk modellering, National Environmental Research Institute, Aarhus University, Aarhus, 1-90 pp., 2010.

Markager, S., Stedmon, C. A. and Sondergaard, M.: Seasonal dynamics and conservative mixing of dissolved organic matter in the temperate eutrophic estuary Horsens Fjord, Estuarine Coastal and Shelf Science, 92, 376-388, 2011.

Olesen, J. E., Carter, T. R., az-Ambrona, C. H., Fronzek, S., Heidmann, T., Hickler, T., Holt, T., Minguez, M. I., Morales, P., Palutikof, J. P., Quemada, M., Ruiz-Ramos, M., Rubaek, G. H., Sau, F., Smith, B. and Sykes, M. T.: Uncertainties in projected impacts of climate change on European agriculture and terrestrial ecosystems based on scenarios from regional climate models, Clim. Change, 81, 123-143, 2007.

Paerl, H.W.: Controlling Eutrophication along the Freshwater – Marine Continuum: Dual nutrient (N and P) Reductions are Essential, Estuar Coast, 32, 593-601, 2009.

Paerl, H. W. and Huisman, J.: Climate change: a catalyst for global expansion of harmful cyanobacterial blooms, Environmental Microbiology Reports, 1, 27-37, 2009.

Postma, D., Boesen, C., Kristiansen, H. and and Larsen, F.: Nitrate reduction in an unconfined sandy aquifer – water chemistry, reduction processes and geochemical modeling, Water Resour Res, 27, 2027-2045, 1991.

Ptacnik, R., Andersen, T. and Tamminen, T.: Performance of the Redfield Ratio and a Family of Nutrient Limitation Indicators as Thresholds for Phytoplankton N vs. P Limitation, Ecosystems, 13, 1201-1214, 2010.

Quile, K., Becker, A., Kirchesch, V. and Schöl, A. and Fisher, H.: Influence of global change on phytoplankton and nutrient cycling in the Elba River. , 11, 405-42., Reg. Environ. Change, 11, 405-442, 2011.

Rockstrom, J., Steffen, W., Noone, K., Persson, A., Chapin, F. S., Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sorlin, S., Snyder, P. K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R. W., Fabry, V. J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P. and Foley, J. A.: A safe operating space for humanity, Nature, 461, 472-475, 2009.

Comment [d56]: Check against text

Sonnenborg, T. O., Hinsby, K., van Roosmalen, L. and and Stisen, S.: Assessment of climate change impacts on the quantity and quality of a coastal catchment using a coupled groundwater-surface water model, Clim. Change, DOI: 10.1007/s10584-011-0367-3, 2011.

Stedmon, C. A., Markager, S., Sondergaard, M., Vang, T., Laubel, A., Borch, N. H. and Windelin, A.: Dissolved organic matter (DOM) export to a temperate estuary: Seasonal variations and implications of land use, Estuar Coast, 29, 388-400, 2006.

Thodsen, H., Kronvang, B., Andersen, H.E., Larsen, S.E., Windolf, J., Jørgensen, T.B., Troldborg, L.: Modeling diffuse nitrogen loadings of ungauged and unmonitored lakes in Denmark: application of an integrated modeling framework. International Journal of River Basin Management, 7(3), 245-257, 2009.

Thorling, L.: Groundwater - status and trends 1989-2006, Geological Survey of Denmark and Greenland, in Danish (English summary), www.grundvandsovervaagning.dk, 1-79 pp., 2007.

Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., Schindler, D., Schlesinger, W. H., Simberloff, D. and Swackhamer, D.: Forecasting agriculturally driven global environmental change, Science, 292, 281-284, 2001.

Timmermann, K., Markager, S. and Gustafsson, K. E.: Streams or open sea? Tracing sources and effects of nutrient loadings in a shallow estuary with a 3D hydrodynamic-ecological model, J. Mar. Syst., 82, 111-121, 2010.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H. and Tilman, G. D.: Human alteration of the global nitrogen cycle: Sources and consequences, Ecol. Appl., 7, 737-750, 1997

Windolf, J., Jeppesen, E., Jensen, J. P. and Kristensen, P.: Modelling of seasonal variation in nitrogen retention and in-lake concentration: A four-year mass balance study in 16 shallow Danish lakes, Biogeochemistry, 33, 25-44, 1996.

Windolf, J., Thodsen, H., Troldborg, L., Larsen, S. E., Bogestrand, J., Ovesen, N. B. and Kronvang, B.: A distributed modelling system for simulation of monthly runoff and nitrogen sources, loads and sinks for ungauged catchments in Denmark, Journal of Environmental Monitoring, 13, 2645-2658, 2011.

Comment [d57]: Not used in text

Figure Legends

Fig.1. Location and delineation of the investigated estuary and catchment, incl. stream gauging stations (triangles) and national monitoring site below farm land (square).

Fig.2. Conceptuel model of the catchment to Horsens estuary with indication of data and nutrient sources. The work process in calculation of threshold values (TVs) for streams and groundwater is indicated. The "DK-QN" model complex (or NLES & DK-QN & DK-model complex) is a combination of an empirical N-leaching model (NLES, Kronvang et al., 2009b), an empirical monthly flow-weighted N-concentration model from diffuse sources (DK-QN, Windolf et al., 2011) and a physically distributed integrated hydrological model (DK-model, Henriksen et al., 2003).

Comment [d58]: General remark. Check for all et al. references as to whether they need to be written in italic font type or not

Fig.3. Measured and simulated total N concentrations at the gauging station on Hansted stream (see Fig.1).

Fig.4. Historical development of freshwater discharge and N and P loadings to Horsens Estuary (upper figures), discharge weighted concentrations in the freshwater outlet to the estuary and annual average concentrations (0-10 m) in the inner part of the estuary (lower figures), 1984 – 2010.

Fig.5. Modeled nitrogen fluxes in the catchment to Horsens Estuary and the net loading to the estuary, annual average for the baseline period 2000-2005.

Fig.6. a) Time series of inorganic nitrogen concentrations (DIN) from May to October (filled circles, response variable, the two open circles indicate outliers from the model, see Fig.7.), average of monthly total N loads from January to October (open squares, x-variable, tons month^{-1}) and wind speed raised to the third (open triangles, x-variable, (m s^{-1})$^{-1}$. b) residual from model.

Fig.7. a) Observed and modeled values for inorganic nitrogen (DIN), average values from May to October from 1985 to 2006. Filled circles are values from 1985 to 2002, used in parameter selection. Open circles are values from 2003 to 2006 omitted and used for validation. + values from 1993 and 1994 are identified as outliers. b) As a, but all values from 1985 to 1992 and 1995 to 2006 are used for estimation of coefficients. Model: DIN (May-October, normalized) = 0.5570 * N-load (January-October, normalized) + 0.52 * Wind^3 (January the year before-October, normalized), R^2=0.7.

Fig.8. a) Relationship between mean concentration and percent of days with limitation for inorganic nitrogen, DIN and b) inorganic phosphorous, DIP. Calculated annually from 1985 to 2006 for Horsens estuary: filled circles (inner part), open circles (outer part), respectively. For DIN the calculations are performed on data from May to October (184 days) and limitation is assumed to occur when DIN < 14 µg L-1. For DIP the period is from March to July (153 days) and limitation is assumed to occur when DIP < 6.2 µg L-1. The vertical dashed lines indicate when limitations occur for 2/3 of the time, and the corresponding concentrations (DIN 21 µg L-1, DIP 7 µg L-1) are considered the target values for good

ecological status of the estuary. The vertical dotted line is the resulting DIN concentration for the outer part of the estuary with an annual N-load of 560 tons yr-1.

Fig.9. Nitrate-N concentrations (mg L^{-1}) in groundwater monitoring wells (latest measurement). Most monitoring wells are located in anaerobic groundwater and therefore contain no nitrate and low dissolved inorganic nitrogen concentrations (DIN).

TABLES

Table 1. Average N and P concentrations in aerobic and anaerobic subsurface waters measured at agricultural monitoring sites (LOOP3 and LOOP4) for the period 2000-2005 compared to average N and P concentrations measured in the general groundwater monitoring program in the catchment to Horsens Estuary for the monitoring period (1989-2009).

Sample / "well" type	N wells[a] / analyses[b]	NO_3-N mg/l	NH_4-N mg/l	DIN mg/l	TN mg/l	TP mg/l
Agricultural monitoring site						
Average 2000-2005						
UZ – suction cups (LOOP3)[c]	-	8.4	-	-	11[d]	0.013
Drains (LOOP4)[c]	-	-	-	-	*12[d]*	*0.050*
Drains / root zone leachate (modelled, this study)	-	-	-	-	15[d]	-
All wells, 1.5 - 5 m (LOOP3)	22 /444	8.5	0.016	8.5	8.5	0.019
Aerobic wells (LOOP3)[e]	20 /414	9.1	0.014	8.1	9.0	0.018
Anaerobic wells (LOOP3)[e]	2 /30	0.052	0.049	0.12	0.23	0.029
Groundwater monitoring						
Average 1989-2009						
All wells with data in period	119 /183	0.25	0.20	0.47	-	0.13
Aerobic wells	7 /12	2.9	0.051	3.4	-	0.16
Anaerobic wells	112 /171	0.068	0.21	0.28	-	0.13

[a]Number of wells, [b]maximum number of analyses, [c]data from *Grant et al*. [2007], [d]flow weighted concentrations, LOOP3 and LOOP4 are monitoring sites, which are located approximately 2 and 100 Km from the investigated catchment in areas with similar clayey soils [e]"aerobic" and "anaerobic" wells are here defined as wells with NO_3-N ≥ 0.25 mg/L and NO_3-N < 0.25 mg/L.

Table 2. Average N and P concentrations in streams and coastal waters, 2000-2005.

Surface water sampling station	DIN mg/L	TN mg/L	PO_4-P mg/L	TP mg/L
Hansted Stream - (FWm /FWs)[a]	4.9 / -	5.6 / 5.5[b]	0.041 / -	0.10 / -
Bygholm Stream - (FWm /FWs)[a]	7.4 / -	8.0 / 6.6[b]	0.072 / -	0.14 / -
Streams ungauged catchm. (FWm /FWs)[a]	- / -	- / 6.2[b]	- / -	- / -
Horsens inner estuary	0.24	0.55	0.013	0.056
Horsens outer estuary	0.14	0.39	0.011	0.046
Belt Sea	0.04	0.25	0.012	0.040

[a]flow weighted, FWm = measured concentration, FWs = simulated concentration,

[b]Measured and simulated stream concentrations include diffuse and point sources

Table 3. Nitrogen and phosphorus sources and loadings to the Horsens Estuary, 2000-2005, [partly from *BLST, 2010*]

	N Tonnes	P Tonnes	N %	P %
Natural background (NB)	179		17	
Agriculture (AGRI)	704	16.2	65	69
Scattered dwellings (SD)	15		1.4	
Industrial plant discharges (IP)	0	0	0	0
Fish farming (freshwater) (FF_{fresh})	0.5	0.07	0.05	0.3
Fish farming (marine) (FF_{marin})	11	1.39	1.0	5.9
Waste Water Treatment Plants (WWTP)	64	1.9	5.9	8.1
Urban stormwater runoff (USR)	15	3.5	1.4	15
Atmospheric deposition on freshwater bodies (A_{fresh})	4.1	0.08	0.4	0.3
Atmospheric deposition on marine waters (A_{marin})	94	0.24	8.7	1.0
Sum all sources	1086	23.4	100	100

Table 4. Coefficients from the empirical models, the maximum observed concentration (µg L^{-1}) in the period 1985 to 2006 (year in brackets) and estimated values with the empirical models and normalized climate: 2001-2005, with target loads for good ecological status, and background loads. Load for nitrogen and phosphorous are given in bracket in metric tons of N or P $year^{-1}$.

	TN 1-12*	DIN 5-10*	TP 1-12*	DIP 3-7*	TN 1-12*	DIN 5-10*	TP 1-12*	DIP 3-7*
	Inner estuary				Outer estuary			
Coefficients µg L^{-1} (tons N or P $year^{-1}$)$^{-1}$	0.20 (N)	0.023 (N)	0.46 (P)	0.20 (P)	0,13 (N)	0.017 (N)	0.33 (P)	0.07 (P)
Maximum obs. values 1985-2006 (µg L^{-1})	836 (1990)	107 (1993)	97 (1986)	21 (1988)	646 (1990)	52 (1993)	58 (1986)	13 (1993)
Estimated values, 2001-2005 (N=1086, P=23.4, Table 3)	567	32	48	8.1	421	14	35	6.9
Estimated values with target loads (N=560, P=13)	462	20	43	6.0	355	5.3	31	6.2
Estimated values with back-ground loads (N=252, P=8.1)	401	12	41	5.0	316	0.1	30	5.9

*The numbers refer to the months over which the average values are calculated.

Table 5. Model calculated annual average N-leaching and flow-weighted N concentrations in root zone water (1 m depth) from agricultural land within the three sub-catchment to the Horsens Estuary.

Agro-hydrological years	Average N-leaching from root-zone on agricultural land ($kg\ ha^{-1}\ yr^{-1}$)	Flow-weighted N-concentration from root zone on agricultural land ($mg\ L^{-1}$)
Hansted Sub-catchment (136 km^2)		
2000	48.1	15.9
2001	85.3	18.7
2002	50.5	15.6
2003	52.8	22.7
2004	73.4	16.9
2005	35.1	22.6
Average	57.5	18.7
Bygholm Sub-catchment (154 km^2)		
2000	48.2	16.5
2001	98.0	18.5
2002	53.3	15.0
2003	55.5	22.1
2004	78.0	17.5
2005	39.0	20.5
Average	62.0	18.4
Ungauged catchment (228 km^2)		
2000	42.8	17.1
2001	73.0	19.8
2002	43.1	16.8
2003	42.6	26.9
2004	63.9	17.8
2005	31.3	23.8
Average	49.5	20.4

Table 6. Modelled N-leaching and gross N-emission from diffuse sources within the catchment to Horsens estuary during the period 2000-2005. The total N-removal in groundwater and surface water is also shown for the same period. Loadings are in metric tons, concentrations in mg L^{-1}. Numbers in parenthesis indicate the percentage of amount leached from the root zone.

Agro-hydrological years	N-leaching from the root zone	Modelled gross N-emission from diffuse sources	N-removal in ground water[a]	N-removal in surface water[b]	Net N-loading[c] to Horsens Estuary	Average stream flow-weighted N-concentration[d] at inlet to estuary
2000	1851	1070	780 (42)	224 (21)	846	5.6
2001	3384	1519	1865 (55)	263 (17)	1256	6.0
2002	1952	1014	937 (48)	205 (20)	809	4.7
2003	1973	793	1180 (60)	189 (24)	605	5.0
2004	2856	1093	1763 (62)	233 (21)	860	5.1
2005	1390	669	721 (52)	168 (25)	501	4.4
Average	2234	1026	1208 (53)	213 (21)	813	5.1

[a]Percentage removed in groundwater is calculated as N-removal divided by N-leaching. [b]Percentage removed in surface water is calculated as N-removal divided by the sum of modelled gross N-loss from diffuse sources and point sources discharges of N (90 t yr^{-1}). [c]Land based loading from diffuse sources (excl. N from atmospheric deposition and sewage outlets).[d]excl. point source contributions

Table 7. Scenario for reductions in total nitrogen and total phosphorus to Horsens Estuary where mitigation measures are only directed at diffuse sources in the catchment. The required reduction is in metric tons, the concentrations are in mg L^{-1}.

	Scenario 1	
	Total N	Total P
Reduction in Diffuse sources	526	10.4
Current stream concentration	6.2	0.15
Stream threshold concentration	2.9	0.084
Current groundwater concentration	15[a] / 0.3[b]	0.018[a] / 0.13[b]
groundwater threshold concentration	6.0	-

[a]aerobic groundwater / [b]anaerobic groundwater

Table 8. Scenario for reductions in total nitrogen and total phosphorus to Horsens Estuary. Reduction targets are in metric tons, concentration in mg L^{-1}.

	Scenario 2	
	Total N	Total P
Total reduction target Horsens Estuary	526	10.4
Reduction in point sources		
Closing of marine fish farm	11	1.39
50% reduction larger point sources	40	2.75
Total	51	4.14
Reduction in atmospheric deposition		
25% reduction atm. deposition	25	-
Remaining reduction target Horsens Estuary	450	5.90
Targeted Mitigation Measures in catchment		
Restored riparian wetlands (300 ha)	60 [a]	3.0 [a]
10 m buffer zones with tree planting along 300 km watercourses [2]	24 [a]	3.0 [b]
Remaining reduction implemented as general mitigation measures on diffuse sources	366	0
Stream threshold concentration	3.1	0.084
Groundwater threshold concentration	9.3	-

Mitigation measures are directed both at point sources and atmospheric deposition from agriculture and targeted as well as general mitigation measures against diffuse sources are utilized. [a]Immediate reduction, [b]Longer term reduction (10-30 years),

Table 9. Current groundwater and stream concentrations and calculated threshold values (TV) for total N and P. The TVs are computed for the two described scenarios (management options) described in the text. All values are in mg L^{-1}.

		Current conc.	TV Scenario 1	TV Scenario 2
Groundwater (aerobic part)	Total N	15[a]	6.0[a]	9.3[a]
	Total P	0.018[b]	-?[c]	-?[c]
Streams	Total N	6.1[d]	2.9	3.1
	Total P	0.15	0.084	0.084

[a] based on the combined use of monitoring and modeling data for the period 2000-2005, [b]Based on monitoring data only, [c]Estimation still not possible - more research is needed, [d]Average of modeled concentrations in the three subcatchments of Horsens estuary.