Authors' interactive comment on BG Discussion paper 9, 8859-8904, 2012 by T. Dalgaard *et al.*: "Farm nitrogen balances in six European agricultural landscapes – a method for farming system assessment, emission hotspot identification, and mitigation measure evaluation".

We thank you very much for the constructive comments submitted by the two referees. In general we agree in these proposals for paper improvements, and our responses are summarized in the following, in blue.

Response to referee #1, Biogeosciences Discuss., 9, C3026–C3030, 2012, by martin.volk@ufz.de. Received and published: 7 August 2012:

This paper presents a method to produce nitrogen balances on a farm level in six European landscapes. Such studies are extremely important with regard to the efficient use of fertilizers and also for generating knowledge bases for nutrient emission mitigation strategies under different conditions. The manuscript is with a few exceptions clearly written and (also with a few exceptions) structure, conditions, management, databases as well as lacks and shortcomings are comprehensively described.
Due to the complexity of the topic, the authors touch several subtopics of which I think that some have to be described more deeply and others can be shortened or left out. Before I will list some suggestions for potential modifications, I would like to congratulate the authors for their great work and efforts that authors put in this study. I suggest moderate revision of the manuscript. Overall, I think this paper would make a very valuable contribution to the journal "Biogiosciences".
General comment: Reduce number of the word "however" ;o)
We thank the referee for his comments. The paper has been revised according to the comments listed below, and the word "however" is deleted as suggested in line 9 and line 22 of the abstract (p. 8861), section 2.4 line 21 (p. 8873), section 3.1 line 10 (p. 8875), section 3.2 line 28 (p. 8875) and in section 4.2 line 9 (p. 8882).

Title:
In my opinion, the title is too long. A potential alternative could be: "Farm nitrogen balances in six European landscapes as indicator for N losses and basis to improving N management."
We agree, and suggest the title is revised to "Farm nitrogen balances in six European landscapes as indicator for nitrogen losses and basis for improved management" (avoiding the use of the abbreviation "N" in the title).

Abstract:
I would start with one sentence about the importance of N studies / "the N problem" for the environment. The next sentence could be on what currently is lacking, then present your method. Currently, there is no explanation about your method in the abstract. I would split the (current) first sentence. Mention that N balances on this scale (or as you have done it) are special and innovative (link to other scales!). Add a sentence on uncertainties (important for future studies to filling research gaps).
We suggest the introduction of the abstract is revised to "Improved management of nitrogen (N) in agriculture is necessary to achieve a sustainable balance between the production of food and other biomass, and the unwanted effects of N on water pollution, greenhouse gas emissions, biodiversity deterioration and human health. To analyse farm N-losses and the complex interactions within farming systems, efficient methods for identifying emissions hotspots and evaluating mitigation measures are therefore needed. The present paper aims is to fill this gap at the farm and landscape scales. Six agricultural landscapes in Poland (PL), the Netherlands (NL), France (FR), Italy (IT), Scotland (UK) and Denmark (DK) were studied, and a common method was developed for undertaking farm inventories and the derivation of farm N balances, N surpluses and for evaluating uncertainty for the 222 farms and 11,440 ha of farmland included in the study."

- line 17: "We conclude..". How can you conclude that here? This sentence stands a bit alone here (after the last section).

We suggest the word "conclude" is changed to "found".

- line 22: "However, no significant...". Short time factor might be one reason, but the other question is if you could cover the different coniditions such as soils, weather, denitrification potentials, location, dimension and hydrology of rivers with your monitoring strategies. I am not sure if your correlations make sense in this regard.

We suggest the sentence is revised to "No significant correlation was found with N measured in surface waters, probably because of spatial and temporal variations in groundwater buffering and biogeochemical reactions affecting N flows from farm to surface waters."

1 Introduction:

page 8863, lines 1 to 21: I think you should give also some examples for modelling / monitoring studies on different scales (see at least some papers in a special issue of Agriculture, Ecosystems & Environment 142 (guest editors Volk & Ewert, 2011) which deal with data and models (and with related problems) on regional to continental scale). There is also (beside many others) some nice work by Hewett et al. (2009) on multiscale management of diffuse pollution (Env. Modell. & Software 24, 74-85), other related studies have been done by Louise Heathwaite, H. Jarvie, I.G. Littlewood or Ullrich and Volk (2010, Env. Monitoring & Assessment 171, 513-527) as well as Volk et al. (2009, Land Use Policy 26, 580-588). Of course you do not have to cite all these studies, but it would be good to give some more background about agriculture- and WFD-(and nitrate directive) related nitrogen measuring, modelling and monitoring on different scales.

We have added references to Volk et al. (2009) in line 27 (p.8862), to Ullrich and Volk (2010) in line 3 (p.8863), and to Hewett et al. (2009) in line 11 (p. 8863).

Hewett, C.J.M., Quinn, P.F., Heathwaite, A.L., Doyle, A., Burke, S., Whitehead, P.G., Lerner, D.N.: A multi-scale framework for strategic management of diffuse pollution. Environ. Model. Software 24, 74-85, 2009.

Ullrich, A., Volk, M.: Influence of different nitrate-N monitoring strategies onload estimation as a base for model calibration and evaluation. Environ. Monit. Assess. 171, 513-527, 2010.

Volk, M., Liersch, S., Schmidt, G.: Towards the implementation of the European Water Framework Directive? Lessons learned from water quality simulations in an agricultural watershed. Land Use Policy 26, 580-588, 2009.

2 Materials and Methods
page 8864, ff
2.1 Study landscapes

- It is not clear to me how you came to the selection /choice of the farms. Could you give an explanation?

In line 16 (p.8864) we have added: "and perspectives for further elaboration of existing studies and data collections (Bienkowski et al., 2009; Hansen et al., 2004; Molenat et al., 2004)", including a new reference to:

Bienkowski, J., Jankowiak, J., Kedziora A and Wasilewski, J.: Possibilities for the Development of Mixed Crop-Beef Farming in the Koscian Region Based on Natural Fodder Resources. Results from the MEA-Scope Case Study in Poland. In: Piorr, A. and Müller, K. (eds.) Rural Landscapes and Agricultural Policies in Europe. P. 269-283. Springer Verlag, Berlin, 2009.

- I think it would be useful to have a bit more information on the river networks (see my previous comment), for instance on location, dimension, hydrology, connection to fields, etc.

In line 4 (of p. 8865) we have added: ", and contributed to larger river networks (Hansen et al., 2004; Bienkowski et al., 2009; Molenat et al., 2004)."

- Perhaps you could shorten the text a bit if you would add a table including information on size, management, live stock, etc., for the farms.

The figures have already been synthesized in Table 3, and in Figures 1 and 2.

2.2 Farm N balance and surplus

- It would be good to have a flow chart of the method including the data sources / data flow. Perhaps you could extend Fig. 3 (add inputs and outputs to the figure) and do another one with the data? I mean you use interviews, assumptions, measurements at different years, statistics, different other databases (which becomes obvious also in chapter 2.3). I would visualize that, which also could help to point out better the gaps that we have to work on in the future (and perhaps to quantify uncertainties). As you know, alone the atmospheric N deposition rates are discussed in a very wide range.

A reference to Section 2.3 was inserted in line 5 (p. 8870) via the sentence "Based on the farm data collection described in Section 2.3". Moreover, a more extensive discussion of N derived from the atmosphere has been added according to the comments of referee #2 (see also below).

2.3 Farm data collection
page 8873, line 2 to 10: What do these "exceptions" mean for your results?

In line 9-10 of p.8873 we added "as elaborated in the discussions section, this was considered not to have significant consequences for the overall results".

2.5 Verification measurements
- Great and important procedure - did you use a common measurement / monitoring design for all you study areas? Nitrate concentrations were measured once a month? This is not much to cover variabilities - see on impact of sampling strategies Ullrich and Volk (2010, Env. Monit. & Assessm.)

In line 11 (p. 8874) we made the following addition: "See also Schelde et al. (2012) and Wohlfart et al. (2012) for more information about the monitoring designs, and Ullrich and Volk (2010), Bende-Michl et al. (2011) for a further discussion of sampling strategies and treatment of the variability in results, or Vogt et al, (2012), for a detailed investigation of N budgets derived from measurements and modelling in the UK landscape.". This is also described in Theobald et al. (2011), which is already referred to.

3.3 Comparison with independent N measurements
line 23 to 28 and following 3 lines on the next page 8876: As mentioned before - you have different locations, soils, weather conditions, denitrifcation potentials, hydrological conditions, management schemes. What monitoring design did you use to investigate these complex relationships in such a short time?

In line 26 (p. 8876) we have added the following new text and references "; additional, more long-term studies of relationships with the local hydrology, soils, weather etc. would be required before conclusions can be drawn (Bende-Michl et al., 2011; Ullrich and Volk 2010; Vogt et al., 2012).".

3.4 N surplus variation and hotspots farms
page 8877
line 5-9, Split / simplify sentence.
line 16: "in Fig. 6"

We have added brackets to the sentence "To explore the N-surplus variation (indicated by the confidence intervals of Fig. 4), and to identify and discuss specific N-surplus hotspots and potential N-mitigation measures in the six landscapes, the farm N-surplus values were plotted against the average farmland N input in the form of synthetic fertilisers, manures, N fixation and deposition from the atmosphere (Fig. 6).", and an "n" in "in" (line 16, p. 8877)

3.5 Example on the effect of N mitigation measures in the Danish landscape
page 8878
This is a very nice example, but it is not clear how it is related to the other studies / farms. I would bring this either section more into context to the other farm studies or remove it. Right now, this section stands a bit alone here. Maybe write a another paper on mitigation strategies based on your studies ;o)

We agree this example could form the background for another paper, but also find it important to include it as an example in the present paper. To emphasize it is an example for one and not all landscapes, the following introduction to section 5.3 (line 13 p 8878) has been added: "A special case study from the Danish landscape is used to illustrate the use of the presented N-balance method for N mitigation measure evaluation."

4 Discussion
page 8879
For me the greatest value of your study is to highlight the hotspots, N balances on the farm level in different landscapes as well as to show the different databases, methods and assumptions used to perform the procedure in these different regions with different conditions. This is a great basis for scale- and region-specific development of measures for N emission reduction. So, for me it is not most important to compare the farms amongst each other but more to point out the specifics of conditions, vulnerabilities, tradeoffs which all is good to show environmental problems but also gaps and lacks in data, methods, etc.
This is a very good point, and we have added the following text in p. 8879, line 19: "Moreover, the study highlights hotspots in the form of critically high N balances at the farm level in different landscapes, and shows the different methods and assumptions used to perform the N balance evaluation procedure in these different regions. This provides a background for region specific development of measures for N emission reduction at both farm and landscape scale."

4.4 Landscape-scale measurements and sustainable farm N management design - add hyphen between "landscape" and "scale"
page 8885-86
- As mention before, I think it is difficult to adjust the monitoring design to the different conditions of all these diofferent landscapes and farms in order to cover all the relevant processes.
- You describe in this chapter only the Danish (nice) example - so what N mitigation strategies would you suggest for the other landscapes of your studies?
A hyphen was added in line 3 of p. 8885.
In line 15 of p. 8886 we have added "The present study only covered results from one year, and did not include a closer investigation of this, but further investigations would be interesting for future landscape studies."

Response to referee #2, Biogeosciences Discuss., 9, C4221–C4224, 2012, by sissel.hansen@bioforsk.no, Received and published: 23 September 2012:

GENERAL COMMENTS
Increasing loads of nitrogen in soil, air and water with the following locally and globally pollution problems do make mitigation strategies urgently needed. To facilitate mitigation, tools are helpful to identify hot spots of nitrogen and possibilities to reduce N-surplus and increase N-efficiency. This paper, focusing on the possibilities and challenges with farm gate N-balances on commercial farms, is therefore an important contribution. Farm gate N-balances is a simple, but rough tool. The paper gives a good overview on the method, and how they used it across a range of European agricultural landscapes. This method has been used for a long time in Denmark. New is that the farm gate N-balances is simultaneously estimated on highly diverse agricultural landscapes, and related to ammonia concentrations in air and nitrate levels in soil and groundwater in the same landscapes.
The paper is interesting, well written and the language is easy to read. However, it could have been more concise if the structure of the paper was stricter. The figures are illustrative and good.
We thank the referee for her comments. Regarding the suggested revisions, with movements and additions of text sections, we think the paper has improved to be more concise.

SPECIFIC COMMENTS
LOCAL FOCUS ONLY? The purpose of the present study is not clear. Is the purpose only to address nitrogen surplus causing local problems, or is globally consequences of Nitrogen surplus also an issue.
This is an important point to be aware of, and make more precise in the paper (see below). In general, N-related greenhouse gas emissions have global consequences, whereas the consequences of nitrates and ammonia losses primarily have local (or regional) consequences.

If the first is the case, N-surplus in the present landscape is what matters, and N-efficiency should only be used as a tool to identify mitigation possibilities. A surplus of nitrogen brought out of the landscape in i.e. dried manure should not be included.
If the latter is the case, N-efficiency, or even better, the N-cost must be included (Bleken et. al. 2005). It is in then important also to include estimates of the nitrogen costs of fodder and dressing imported to the farm. As there are records of what is imported in the present investigation, this should be possible using literature values. In page 8882 the authors state "the systems with the highest N inputs are also those which are most N-efficient". They found the highest N-efficiency in the French agricultural landscape and the lowest in the Scottish. However, the French farms had the highest fodder import and the Scottish the lowest (Fig.4). If the N-costs of imports were included, I assume the conclusions would have been different. In calculation of global consequences of nitrogen use, manure nitrogen brought out of the landscape should be a part of the surplus and not calculated as N in products.
We have added the following text section to the discussion (line 16, p. 8882): "This study primarily focuses on local (farm or landscape level) effects of N-losses in the form of ammonia or nitrate, but also the more global consequences in the form of N-related greenhouse gas emissions are important. For the evaluation of such global consequences of N-surplus, N-efficiency, or N costs in the form of N losses associated with the production of the fodder and manure imported to the farm must be included (Bleken et al., 2005). In the present study, the highest N-efficiency was found in FR and the lowest in UK. However, the French farms had the highest fodder import and the Scottish the lowest (Fig. 4). Consequently, if the N-costs associated with fodder imports were included, the conclusions might have been reverse. Similarly, in the calculation of global consequences of N use, manure N removed from a landscape aught instead be a part of the surplus and not calculated as N in products."

N-FIXATION: In low input systems the quality of assessments of nitrogen fixation will have a large impact on the accuracy of the N-balance. One level for nitrogen fixed via pulses (100 kg N ha$_{-1}$yr$_{-1}$) and two N levels for legumes in grassland (150 or 20 kg N ha$_{-1}$yr$_{-1}$) are very inaccurate, as stated by the authors. In such a rough study it is not easy to get good estimates of the roughage yields or the proportion of legumes, neither in the harvested roughage nor in the grazed areas. However, this large uncertainty should be accounted for when estimated surpluses and N-efficiency are discussed.
In line 24 (p. 8880) we have added: "Especially in low input systems the accuracy of estimates of N fixation will have a large impact on the accuracy of the N-balance. As stated, one estimate for N fixed via pulses (100 kg N ha-1yr-1) and two N levels for legumes in grassland (150 or 20 kg N ha-1 yr-1) are very coarse estimates covering a considerably larger variation. In a study that relies on farmer interviews, it is not easy to get good estimates of the roughage yields or the proportion of legumes, neither in the harvested roughage nor in the grazed areas. Consequently, this large uncertainty should be considered when assessing estimated surpluses and N-efficiencies."

STANDARD VALUES FOR N: The option to exchange the standard N-content with local values was only used in a few cases by the local partners. The authors stated "From this we concluded that there is no reason to expect a systematic over- or underestimation from using the current method,..". I would like some more arguments for this statement. The largest fodder import is in France, the Netherlands and Poland. Therefore inaccuracy in N content from fodder import would

have the largest impact in these three cases. For these agricultural landscapes it should be stated if local values were used. If not, it should be stated why. The same should be done if large amounts of manure are exported or imported.

We have inserted the following text p. 8880 line 11: "However, when assessing the results it must be remembered that the largest fodder imports were in landscapes of FR, NL and PL. Therefore any inaccuracy in the import of N in fodder would have had the largest impact in these three cases. The same argument holds for the large amounts of manure which are exported from or imported to the landscapes."

UNCERTAINTY: Uncertainty in the estimates is a challenge when farm gate balances are used to identify hot spots of nitrogen surplus and efficiency. To facilitate that other studies can use the method developed in the present study, there should be more discussion on the uncertainty of the results and how to judge the uncertainty. To better show the estimated uncertainty the illustrative fig. 4 may be used. The arrows showing the estimated surplus for each agricultural landscape can be supplemented with for instance a grey colour in the arrows showing the estimated interval of uncertainty.

Good idea. However, we think this is already included with the ± figures in Fig. 4, and adding gray shaded arrows would make the figure less easy to view and read.

In addition to each value of net fodder, net dressing etc. an uncertainty is given as (±) in the present manuscript. However, it is not explained what this is. It should be stated if this is only variation (standard deviation) caused by differences in the farms within the landscape areas, or if also uncertainties in the estimates are included.

Good point. We suggest to state in line 9 of p. 8871: ", and 95% confidence intervals (based on standard deviations) caused by differences in the farms within the landscape areas are included"

STRUCTURE: The agricultural landscapes are described in M&M, in results and in discussion. To make this shorter and easier to read, the descriptions should be concentrated on the information needed to understand the differences among the systems. I would have preferred to have all this information gathered in M&M, and to refer to this later in the manuscript. All methods should be placed under M&M including definition of N-efficiency.

An equation (2) defining N-efficiency was added to the Materials and Methods section p. 8871 line 8: "In addition, the N-efficiency is defined from the values of Fig. 3 as net N output in products sold, divided by net N inputs purchased by the farmer:: N-efficiency= ([o1+o2-i5] x ([i1+i2-o4]+[i3+i4-o3])-1   (2)", and line 4-5 p. 8882 was revised to "If N efficiency in accordance with equation (2) is defined as net N output in products sold divided by net N inputs purchased by the farmer".

TECHNICAL CORRECTIONS
Page 8870. The term meat is used also for eggs and wool and I assume also live cattle, but it is not stated. I suggest using the term animal products instead and defining what is included.

The words "and other animal produce (incl. meat, live animals, eggs and wool)" must be added to line 6, p.8870.

Table 2. This table is confusing. I suggest sorting the manure according to animal species.

We agree the original order of table was confusing have revised the table in alphabetical order (i.e. the first line with "pig slurry (fattening pigs)" is moved down under "p").

COMMENTS TO THE COMMENTS FROM M. Volk: I do not disagree with any of the comments. Following the suggested ideas will improve the manuscript.

REFERENCE: Bleken, M.A., Steinshamn, H., Hansen, S., 2005. High nitrogen costs of dairy production in Europe: Worsened by intensification. Ambio 34, 598-606.

The reference to Bleken et al. (2005) has been added together with updated references for a number of the other references (for example Schelde et al., 2012 and Christen et al., 2012).