

Authors reply to referees comments on BGD manuscript:

"Greenhouse gas emissions from the grassy outdoor run of organic broilers" by Meda et al.

Introduction

We thank both referees for their largely constructive comments and suggestions. A common view of both referees is that i) the original manuscript is rather too long and ii) the section (4.3) on the full greenhouse gas (GHG) budget is a weaker part of the paper. Both referees nonetheless agree that the data are original, that the paper "contains valuable information worth publication" and that the "approach is commendable, as are the efforts to determine overall fluxes using different methods".

We agree that the strength of the paper resides in the provision of annual estimates of soil fluxes of N₂O, CH₄ and CO₂, for a heretofore little studied system and using a combination of temporal and spatial gap-filling methods to approach the uncertainty therein, rather than in the proposed full analysis (Section 4.3) of the GHG budget and carbon (C) sequestration potential of the system. We also agree that the paper can be reduced in length while maintaining its core messages and overall quality.

One obvious place to reduce the word count is Section 4.3, which was devoted to the net GHG budget of the paddock, but which, admittedly, was based on a healthy dose of speculation as well as hard facts. However, because we believe in the investigative and innovative value of this analysis, and also because we argue that over such an outdoor production system (as opposed to indoor poultry units with no carbon sink) there might exist compensating effects between the different GHGs, an attempt should be made to integrate the global warming potentials of CO₂, N₂O and CH₄ into one overall indicator. We thus propose to shift most of the methodology and reasoning behind the net GHG budget calculation from section 4.3 to a Supplementary Material document, which will be appended to the revised version of the paper. Section 4.3 will be kept to a reduced account of the main results and their uncertainties, and the section will be renamed "*4.3 Towards a net GHG budget of the outdoor run*". The accompanying Tables 5 and 6 will also be moved to Supplementary Material, as will the associated references. Overall, this should reduce the word count of the main paper by approximately 1500 words.

We are confident that these changes will result in a streamlined version of the article, with greater readability towards the end. Detailed replies to the referees comments are provided hereafter, with [our replies highlighted in light blue fonts](#), and *proposed changes to the article text in italics*.

Reply to Anonymous Referee 1

Specific comments: Most importantly, the MS is much too long. Not counting References, Tables and Figures it already has c. 10500 words, with an overall count > 16000. On top of this there are 13 Tables and Figures! Half of that extent should do. The length comes from two sources: First, starting with Mat & Meth, there is both too much unnecessary detail (non-essential information) and the wording is too elaborate. The reader deserves a more condensed account. The second cause touches the scientific content. For example, the Results section 3.3 contains a 1625 words section on temporal gap-filling functions alone. I suggest that a few essential lines of this would suit the discussion well, if the topic of the paper should really be the GHG balance and not a technical note. The Discussion section 4.3 (> 1700 words) contains many explanations and calculations that are completely Mat & Meth and Results material.

As indicated above, we will reduce the length of the paper by shifting most of Section 4.3 and Tables 5 and 6 to Supplementary Material. We do not otherwise believe that Materials and Methods are overly detailed nor should they be reduced. One important aspect of this work lies in the recognition that in outdoor poultry systems the spatial distribution of bird droppings will not be random but highly organised, due to preferential visiting areas within the paddock. This led us to a design of the measurement system and of the experimental approach that was the best compromise between temporal and spatial variability, technical and financial feasibility and overall representativeness of flux measurements. We therefore believe that a comprehensive description of M&M is essential for the reader interested in flux methodology adapted to such systems.

Similarly, the description of the gap-filling methods (Section 3.3) is equally important and does not turn the paper into "a technical note". It is our view that many papers, especially those concerned with N₂O emissions, provide too few details on the methodology used to scale up to annual estimates from discrete (discontinuous) time series. Our section on gap-filling has two objectives: 1- to understand and quantify the control of fluxes by environmental macro-drivers (T, SWC / WFPS), which then enables 2- the simulation of fluxes for periods not covered by measurements, and hence the provision of the annual integral. These objectives are both scientific and technical, in the sense that the diagnostic numerical analysis of the flux data (e.g. multivariate analysis) requires an understanding of the underpinning biology, and then leads to an operational tool for gap-filling. (in CO₂ flux research, there are entire scientific papers dedicated to gap-filling methods and strategies, e.g. Falge et al., 2001, *Ag. For. Met.* 107, 43-69; Reichstein et al., 2005, *Glob. Ch. Biol.*, 11, 1424-1439)

Also, I have a hard time with the CO₂ part of the MS. The study concludes that the chicken run is a substantial short term sink for GHG. This highly important finding is based on a large ecosystem C gain through assimilation (GPP). Ironically, GPP was not measured in this work. Instead the positive CO₂ GHG balance is derived from an estimate of chicken droppings and a short-term change of the soil organic carbon (SOC) content of c. +380 g C m⁻² yr⁻¹. This is certainly < 3% of the total SOC. Despite the fact that it is very difficult to measure such small changes in SOC significantly different from zero, there is no information in Mat & Meth or in Results to show how this very important bit of information was gained?! Compared to the ca. 12 kg C m⁻² the soil may have accumulated in the last 10000 years, a rate of 0.4 kg C in 1 year demands an explanation. Thus, it takes a good argument why the authors assume a SOC balance other than zero. If that much more conservative approach was chosen instead, the NGHGE would be about zero, which was a surprising result, too. For the above reasons I suggest to separate much more clearly between reliable data from the Results section and some more speculative numbers, sparingly used in Discussion, to put measured data into

perspective.

We agree with Referee 1 (and also with Referee 2 -see below) that the full GHG balance analysis was based to some extent on speculative numbers and data from the literature, and that much information was missing to make a robust assessment of the net GHG budget of the paddock. We have already alluded to the large uncertainties in these estimates in the paper (p11554, l8-9). We nonetheless believe that the added value of this analysis is to show the likely magnitudes of fluxes in an outdoor animal production system, for which there are no published data, and to point to knowledge gaps and research needs. Despite the very large uncertainties in the net GHG estimates, the potential of the outdoor grassland for C sequestration should not be discarded, especially since we are effectively dealing with a fertilised system, the productivity of which may be enhanced (fertilisation) or suppressed (trampling) by the animals. Thus the ultimate objective of providing a full GHG budget is fully justified, even if key elements of the budget (GPP, leaching, C imports) were not actually measured in this paper.

We will accordingly rename Section 4.3 "*Towards a net GHG budget of the outdoor run*", which will be reduced to a short discussion of the main results and their uncertainties. The more detailed analysis previously shown in Section 4.3 will be shown together with the methodology in Supplementary Material.

Still, we are confident that the measured changes in SOC are highly significant. The data indicate a short-term increase of SOC of ca +380 g C m⁻² yr⁻¹, and the initial stock (averaged over the paddock) was 3450 g C m⁻², for the 0-15cm topsoil layer, which translates into an increase of ca +11%/yr for this layer. Referee 1 suggests an increase of +3% for the whole soil depth, but this number is not substantiated since the C stock for the lower layers is not known. Topsoil at this site is actually fairly shallow. Also, it should not be overlooked that this site was formerly used as a crop rotation (see M&M), in which C stocks are generally much lower than in permanent grasslands, and thus the increase may partly be explained by the transition from arable to grassland.

We refute the Referee's assertion that "there is no information in Mat & Meth or in Results to show how this very important bit of information was gained". Section 2.8 "Temporal and spatial variations in soil nutrient (P, N, C) stocks" does provide details of the experimental method (see p11537, l10 to l25).

We do acknowledge however that the results for soil nutrient (C, N, P) stocks were not clearly nor sufficiently reported before the Discussion. We thus propose to add a short paragraph in Results, "*3.5 Magnitude and significance of temporal soil nutrient changes*", providing the main results and the uncertainties associated with the measurements and upscaling of the C, N and P stocks. This will draw upon data from a new table (Table S1) supplied as Supplementary material.

We propose the following changes to the text:

"3.5 Magnitude and significance of temporal soil nutrient changes

Soil nutrient stocks for C, N and P at the paddock scale all increased significantly from start to end of the experiment, as shown in Table S1 (Supplementary material). The stocks were determined for the 0-15 cm horizon only, by means of stratified sampling for the 0-5 cm and 5-15 cm layers separately, and the data in Table S1 are therefore lower bound estimates of stock changes for the whole soil depth. The paddock-scale stock increments for the 0-15 cm horizon were +377 gC m⁻² yr⁻¹ (+11% yr⁻¹), +25gN m⁻² yr⁻¹ (+7% yr⁻¹) and + 9 gP m⁻² yr⁻¹ (+20% yr⁻¹). Paired-sample t-test statistics on the 25 paired composite samples (initial vs final) were used since the five soil samples used for each composite sample were extracted from the very same locations (+/- 10 cm) at the start and end of the experiment. The temporal stock

increments were thus statistically highly significant ($P < 0.01$) for all three nutrients.

The main source of uncertainty in the calculation of soil carbon stocks was the evaluation of the fraction of coarse stony elements at the paddock scale, which was of the order of 15% in the 0-15 cm layer. We tested the error propagation from the coarse element fraction to the calculated soil elemental stock, and found that hypothetical random errors for the coarse fraction of 5%, 10% and 30% would translate into an error (of opposite sign) of 1%, 2% and 5%, approximately, for the soil stock. Given the magnitude of observed changes in C, P and N stocks (~10-20% per year), and given that the relevant error is that of the stock change (or difference between two measurements), not that of the absolute stock value at a given time, this study concludes that very significant C, N and P stock changes took place during the experiment."

Table S1 for Supplementary material:

"Table S1. Temporal variations between initial (March 2009) and final (December 2010) C, N and P stocks in the topsoil (0-15 cm) layer, for the composite samples from 25 equally-sized (94 m²) sampling squares of the paddock."

Soil sampling Square Id	Mean distance to house door m	C stock per sampling square				N stock per sampling square				P stock per sampling square			
		Initial kg	Final kg	Variation kg (final-initial)	Variation g/m ² /yr	Initial kg	Final kg	Variation kg (final-initial)	Variation g/m ² /yr	Initial kg	Final kg	Variation kg (final-initial)	Variation g/m ² /yr
1	19	349	340	-9	-56	35	33	-1	-9	4.25	6.30	2	12
2	11	352	371	19	113	31	36	4	27	4.04	7.76	4	23
3	5	301	380	79	480	32	37	6	36	4.09	10.18	6	37
4	11	339	366	28	167	36	36	-1	-3	4.82	7.22	2	15
5	19	341	458	117	710	37	44	7	43	4.14	6.90	3	17
6	24	300	345	45	273	31	34	2	15	3.74	5.61	2	11
7	18	298	327	29	176	31	31	1	3	4.93	5.50	1	3
8	15	317	383	66	403	34	38	4	25	7.90	11.04	3	19
9	18	372	437	64	391	37	43	6	36	7.31	9.89	3	16
10	24	374	382	8	46	38	37	-1	-4	4.00	5.55	2	9
11	31	288	381	94	569	30	36	6	37	3.13	4.07	1	6
12	27	332	414	82	497	34	39	5	30	4.43	5.41	1	6
13	25	318	416	98	595	34	40	6	34	5.62	6.68	1	6
14	27	346	415	69	418	38	41	3	17	5.66	5.65	0	0
15	31	358	459	101	611	39	45	6	37	3.71	5.31	2	10
16	40	279	377	98	594	28	35	7	41	2.44	2.74	0	2
17	36	293	329	37	222	30	31	1	7	2.84	3.09	0	2
18	35	304	383	79	480	31	37	6	34	2.88	3.67	1	5
19	36	322	392	70	423	33	37	4	26	3.01	3.46	0	3
20	40	382	407	25	150	36	40	4	27	3.17	3.90	1	4
21	49	298	341	43	259	29	33	4	23	2.10	2.04	0	0
22	46	262	363	101	611	27	35	9	52	2.37	2.97	1	4
23	45	321	388	67	408	34	38	4	23	3.02	3.44	0	3
24	46	321	391	69	421	34	40	7	40	2.95	3.56	1	4
25	49	338	412	74	452	37	41	4	25	3.45	3.67	0	1
Sum (paddock)		8104	9655	1551	377	833	936	103	25	100	136	36	9

Section 4.3 will be thus reduced (and the rest be exported to Supplementary material):

"4.3 Towards a net GHG budget of the outdoor run

Emission factors are needed from an inventory viewpoint (IPCC), but to assess the whole climate change impact of a production system, a total net GHG exchange (NGHGE), taking into account all global warming potential (GWP) contributions by CO₂, N₂O and CH₄ fluxes is required (Soussana et al., 2007). In this study we determined the net N₂O and CH₄ fluxes, but for CO₂ only ecosystem respiration (R_{eco}) was measured; for a full GHG balance at the paddock scale, the net ecosystem exchange (NEE, measurable by eddy covariance, see e.g. Ammann et al., 2007), defined as gross primary production (GPP) minus Reco, would be required. In addition, CO₂ losses by animal respiration ($R_{broilers}$) would add to emissions by plants and soil, as would losses by soil dissolved organic and inorganic carbon leaching, as well as exports by animals as they consume worms and other forms of terrestrial organic materials. C imports by animal excreta on the other hand contribute very significantly to an increase in soil organic matter in this system.

In the absence of measurements for many of these terms, we derived a tentative net GHG budget by inferring NEE from the observed and significant change in soil organic carbon ($+377 \text{ gC m}^{-2} \text{ s}^{-1}$, see Section 3.5) and by making various assumptions concerning C imports, C exports and C leaching. The details of the methodology and reference values from the literature used for this calculation are provided in Supplementary Material. The NEE estimate for the grassland + broiler system (NEE_{G+B}) thus obtained was $-1163 \text{ kgC ha}^{-1} \text{ yr}^{-1}$, implying that the whole ecosystem (soil + vegetation + animals) would have been a net C sink, which was offset by roughly one third by N_2O emissions with a negligible impact by CH_4 oxidation. Overall, the NGHGE of $-2561 \text{ kgCO}_2 \text{ eq. ha}^{-1} \text{ yr}^{-1}$ would suggest that the outdoor part of the rearing system behaves as a net sink of greenhouse gases, but it must be stressed that uncertainties are very large and plot-scale measurements of all component terms are required to provide a more reliable and robust estimate. In addition, much of the short-term (~ 2 -year) increase in soil organic carbon was likely contributed by deposited bird excreta, the longer term fate of which is uncertain, and which therefore does not necessarily represent a potential for C sequestration.”

It seems there were no replicate gas exchange measurements per frame/site. No statistical method for data evaluation is quoted. Concentration changes of GHG in the static chambers were measured four times in 30 minutes. That is a very long period, particularly for CO_2 that may have similar concentrations in the chamber and in the soil at the end of the measurement period. For that reason it is necessary to use the linearity of the regression used here to assure the quality of the measurement. If the linearity is not close to perfect ($> 90\%$), then the measurement is of no value because the chamber has altered (suppressed) the GHG efflux. This additional information is required in section ‘2.3 Flux measurement technique’.

We agree that the theoretical shape of the concentration change over time in the head space is not linear, since from Fick’s law (flux = diffusion coefficient * concentration gradient) we know that the accumulation of CO_2 in the chamber leads to a decrease of the concentration gradient between soil and chamber, and thus the flux should, in theory, be self-limiting. Thus in theory a linear regression may underestimate fluxes significantly if the change in concentration over the time of chamber closure is significant compared with the soil-atmosphere gradient. Although we did not say so explicitly in the manuscript, we assumed that during the 30 minutes over which the 4 samples were taken, the data could be considered linear and that the curvature would be negligible. The reason why we did not choose non-linear functions to calculate fluxes was that the number of gas samples per chamber was only 4 and that measurement errors in individual gas concentration measurements could have a large bearing on a non-linear fit. The linear fit was chosen as a robust, conservative option.

Referee 1 argues that “ CO_2 that may have similar concentrations in the chamber and in the soil at the end of the measurement period”. We consider this a rather unlikely and exceptional occurrence. Consider the CO_2 concentration measured at the end of each chamber measurement, i.e. the 4th sample (CO_2end). Over the entire measurement period, i.e. roughly 300 chamber flux measurements, the median CO_2end was 503ppm, the 75th percentile was 644 and the 95th percentile was 957 ppm. By comparison, continuous soil pore space concentration measurements under permanent grassland (Flechard et al., 2007, Agricultural and Forest Meteorology 142, 66–84) have shown that pCO_2 in soil open pores at -5 cm depth was typically several thousand ppm and frequently comfortably over 10000 ppm in wet conditions, but still typically 2000 ppm in dry conditions.

Nevertheless, we have re-analysed the entire concentration/flux database to assess the potential effect of non-linearity on trace gas fluxes. We fitted a quadratic function ($f = p_2 * t^2 + p_1 * t + p_0$, with t in seconds) to each chamber concentration evolution, for CO_2 , N_2O and CH_4 , with the first-order term (p_1) being equal to the slope of the tangent to the fitted curve (df/dt) at $t=0$ and thus proportional to the chamber flux without feed-back, and the sign and magnitude

of the second-order term (p_2) indicating the shape and degree of curvature.

The results are shown in the following figure (Fig. 1), with the y-axis representing the flux evaluated by quadratic fitting (QuadFlux) versus on the x-axis the flux by linear regression (LinFlux). These data apparently indicate a tendency for linear regression fits to provide lower flux estimates than quadratic fits. The effect appears to be of the order of 26% for CO₂, 16% for N₂O and 29% for CH₄, as indicated by the slope of the linear regression of QuadFlux versus LinFlux. However, for CH₄ and N₂O at least, these slopes are largely driven by large outlying values, while for lower fluxes there is much scatter and no clear pattern.

Figure 2 shows the frequency distribution of the 2nd-order term of the fitted quadratic functions for all three gases. Negative p_2 values (concave fits) indicate quadratic fits where a slowing down of emission over time occurred (chamber feed-back as expected), while positive p_2 values (convex fits) indicate cases when fluxes at the end of the chamber measurement were higher than at $t=0$ (contrary to expectation). For all gases the p_2 parameter was well distributed on both sides of 0, whereas chamber feed-back theory would have all p_2 values negative. Because of this balanced distribution of p_2 around 0, we believe that for the majority of cases the fitted non-linearity is a result of errors in individual concentration measurements, driving random scatter in the 4 values used for the regression. There likely were also cases with a clearly significant non-linearity, though these were in our opinion isolated cases with little bearing on annual fluxes.

To address these issues in the paper, we will add the following sentence at the end of Section 2.3 (“Flux measurement technique”):

“...the slope of the linear regression of gas concentration vs. time was used. It should be recognised that in theory the temporal evolution of concentration inside the closed chamber is curvilinear, owing to a gradual reduction over time of the soil/headspace concentration gradient and hence of the flux. However we did not fit non-linear functions because only four concentration points (every 10 minutes during a half-hour) were available to constrain non-linear models, which, accounting for the uncertainty in individual concentration measurements, may lead to unstable solutions (e.g. Levy et al., 2011, EJSS 62, 811–821). The linear regression slope was deemed a more robust option, and the GHG flux data presented hereafter should be viewed as conservative estimates.”

“Levy et al., 2011, Quantification of uncertainty in trace gas fluxes measured by the static chamber method, EJSS 62, 811–821” will be added to the reference list.

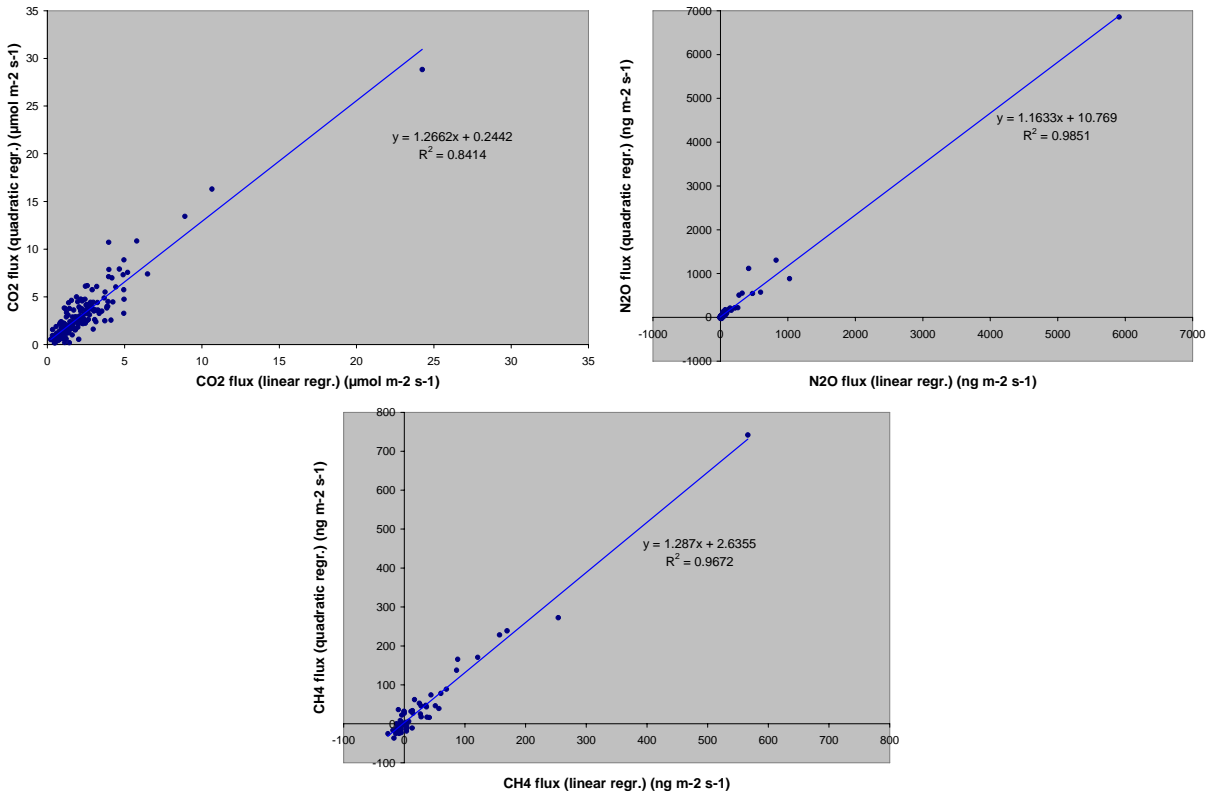


Figure 1. Comparison of fluxes calculated using a quadratic fit versus linear regression, for CO₂, N₂O and CH₄.

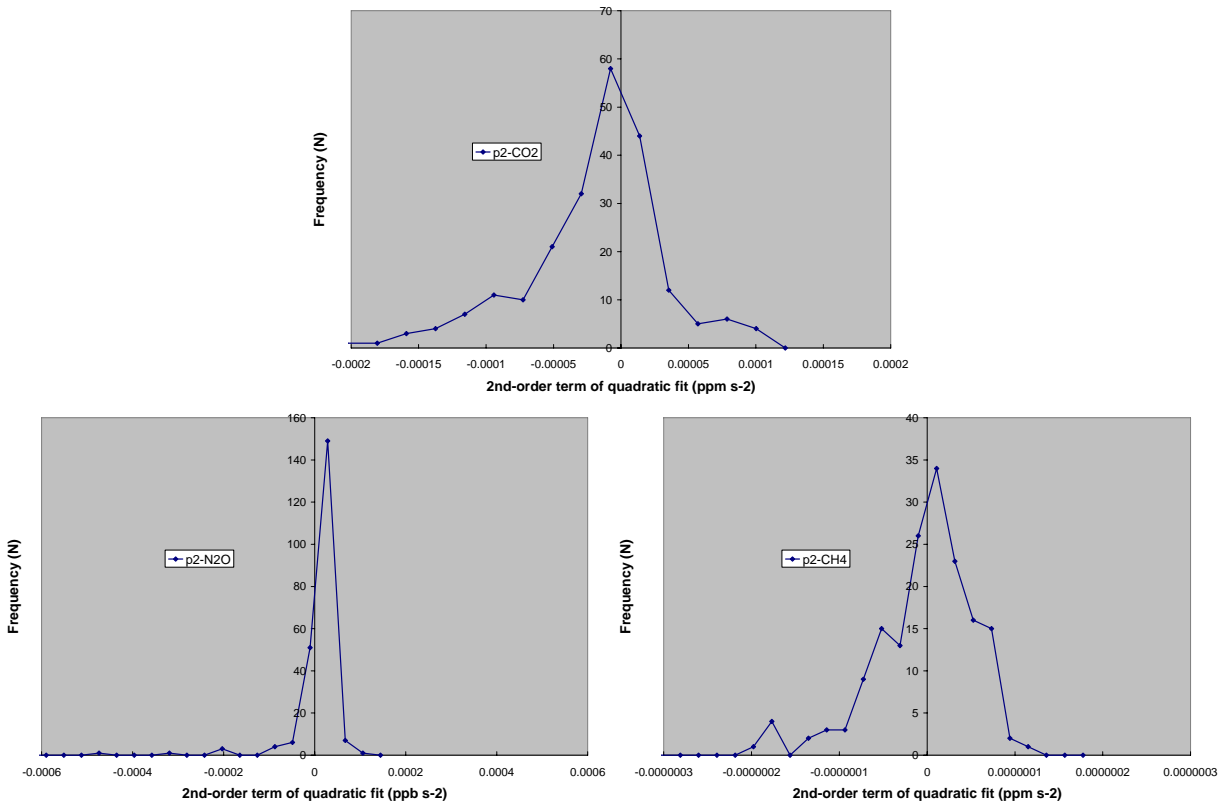


Figure 2. Frequency distributions of the second-order term (p_2) of the fitted quadratic functions for CO₂, N₂O and CH₄.

Publications 'in preparation' should not be in the reference list.

This reference was changed to the PhD thesis by B. Meda published in 2011.

The authors mention the extremely high variability of fluxes between measurement days. I find that interesting, too. But no word is lost in Discussion on apparent inconsistencies that challenge the quality of the measurements. E.g. the maximum CO₂ flux (3.1 $\mu\text{mol m}^{-2} \text{s}^{-1}$) in the WS batch is found at max. soil T (13.9°C) and 100% WFSP. Either WFSP is overestimated or ecosystem respiration at c. 3 $\mu\text{mol CO}_2$ flux does not exhaust soil O₂ availability. The similarly large max. CO₂ flux in the SA batch (3.3 $\mu\text{mol m}^{-2} \text{s}^{-1}$) is found at low 16.5°C and only 38% WFSP. But five of the other eight measurement dates of SA batch have lower CO₂ fluxes despite higher soil temperatures and higher water availability. Similar situation with N₂O: the largest flux is five times larger than the second largest flux. Why? Please explain.

The apparent inconsistency of CO₂ fluxes between different measurement days may, in part, be attributed to measurement errors in soil water content and water-filled pore space, or rather, to the spatial heterogeneity of SWC. We made no secret of the fact that SWC was measured outside the enclosure, that there was a slope in the terrain and therefore heterogeneous conditions for soil moisture. We mention (p 11542, I17) that we are faced with "the challenge of interpreting GHG fluxes mechanistically on the basis of low resolution environmental data". It is very likely that on the day of SA batch when CO₂ flux was 3.3 $\mu\text{mol/m}^2/\text{s}$, WFSP was higher at the soil surface near the broiler house, due to recent rainfall, than the tabulated value (Table 3) of 38%, which was the value measured at -5cm depth on level terrain outside the enclosure. This argument is developed in depth for CO₂ and N₂O on p 11542, and also on p11550, I26 onwards. We acknowledge that a network of TDR probes inside the paddock would have provided us with better spatial information, with which to interpret the GHG fluxes. Unfortunately, in practice the idea was decided against due to the tendency of foraging chickens to dig and scratch the soil surface, particularly in places with slightly protruding unidentified objects such as cables, probes, etc.

Further, the variability in fluxes is also to a large extent attributable to the availability at the paddock scale of freshly deposited C and N in bird droppings. This, in turn, may be dependent on meteorological conditions during the few days or even weeks preceding measurements, as weather affected both the rate of animal outing and their spatial distribution inside the paddock, as already laid out in the paper. Thus a variability in GHG fluxes between days is expected that cannot be explained by meteorological and soil conditions on the day of measurements. For example concentrations of mineral N in topsoil could be useful to explain a larger part of the variability, but these measurements were not made.

Not a single error value on 47 pages of MS. Please change.

The main objective of this paper was to provide an estimate of GHG fluxes, and their uncertainty, from the outdoor run of free-ranging broilers. We did so by combining different methods for both spatial and temporal integration of discrete measurements in space and time, and we have provided numbers and statistics for the different methods used (see Tables 2, 3 and 4). In our view the spread of values provided by the different methods is the best available indicator of the overall uncertainty. For example, for N₂O we provide in Table 4 three estimates of the annual flux (4.8, 3.6 and 2.7 kg N /ha/yr), while in the abstract we gave a "best estimate" of 3 +/- 1 kg N /ha/yr. The paper shows that, depending on the gap-filling methods that are used, very different annual estimates may be obtained, and it also makes it clear that

the uncertainty in such estimates is of the order of 30-50%. For our results we have consistently provided ranges of values, which is just as informative as the value +/-error suggested by Referee 1.

Concerning GHG concentrations and flux determination, we will add to Section 2.3 a few elements concerning the precision of trace gas concentrations by GC and the flux detection limit (lower detectable flux) of the system. We propose to insert the following text on p11534, just before Section 2.4:

“The lower detectable fluxes (LDF) for CO₂, N₂O and CH₄ were determined on the basis of the precision of GHG concentrations measured by GC. Values of LDF were calculated following Flechard et al. (2005) as $LDF=H\Delta C_{noise}/\Delta t_{chb}$, where ΔC_{noise} is the noise or lower detectable concentration difference, and Δt_{chb} is chamber closure duration (30 minutes). The noise levels for the three gases were defined as twice the standard deviation of successive concentration measurements of gas standards during calibrations, and determined in our GC system as ΔC_{noise} (CO₂) = 11 ppm at 600 ppm (precision ~2%), ΔC_{noise} (N₂O) = 21 ppb at 475 ppb (precision ~4%), and ΔC_{noise} (CH₄) = 0.06 ppm at 2.05 ppm (precision ~3%). The resulting LDF values (using H=0.3m) were 0.1 $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$, 6.7 ng N₂O $\text{m}^{-2} \text{ s}^{-1}$ and 7.1 ng CH₄ $\text{m}^{-2} \text{ s}^{-1}$.*

Reply to Anonymous Referee 2

This study presents, apparently for the first time, flux estimates for CH₄, N₂O and Reco, and a full GHG balance calculation, of an outdoor area used for poultry production. On the basis of 19 flux measurement campaigns with static enclosures, conducted mainly in connection with two production batches under winter-spring and summer-autumn conditions, the authors calculate annual fluxes of CH₄, N₂O and Reco by different methods, and the full GHG balance is estimated using production data and informations from a variety of sources. The system studied includes a house where the poultry would spend part of the time, but also an upward slope away from the house. The experimental approach to cover this spatial and temporal variability was a stratified distribution of sampling points and spatial integration by geostatistical methods. This approach is commendable, as are the efforts to determine overall fluxes using different methods. The study does have limitations in the sampling strategy and collection of supporting data, and there is certainly scope for further improvements of flux estimates in future studies. Still, the work provides a comprehensive analysis of a little known aspect of poultry production and could be published following proper revision.

Specific comments: p.11534, l. 2 - Please specify sampling time here. It is stated on p. 11545 to be afternoon, which certainly is not the optimal representation of daily mean temperature, and this bias should be clear to the reader (and potential consequences for flux estimates discussed).

The sampling time varied little and generally started around 14:00 local time (GMT+2 in summer, GMT+1 in winter), and sampling lasted between 1.5 and 2 hours. We will add this piece of information to Section 2.3 “Flux measurement technique” by inserting the following sentence p11534, l3, before “At the start of each flux measurement...” :

“ On each day of flux measurements, chamber sampling was done in the early to mid-afternoon i.e. between 14:00 to 16:00.”

It is true that soil temperature was close to its daily maximum at the end of sampling. The manuscript already states this very clearly on p11545, l8 (“ fluxes were always measured in the afternoon, when soil temperatures and microbial activity were highest.”). However, the gap-filling technique does in a way correct for this bias by accounting for the temperature effect through the Lloyd and Taylor function (Eq. 7), so that simulated night-time fluxes will be reduced accordingly and the annual integral should not suffer too high a bias.

p. 11534, l. 13 - It is relevant to comment on the occurrence of any non-linearity in trace gas accumulation. Poultry manure excreted would be deposited at the surface, indicating that sources of N₂O and CO₂ would also be highest near the surface which theoretically should give a high potential for chamber-feedback.

The non-linearity issue and potential for chamber feed-back were addressed in our reply to Referee 1 – please see above.

p. 11534, l. 15 - How well did the data on soil WFPS, which were based on TDR measurements at 5 cm soil depth outside the fenced in paddock, represent WFPS inside the paddock? This is a relevant concern considering the variation in vegetation, animal traffic and elevation. It would have been feasible to install TDR probes near the sampling points for manual verification of soil moisture trends during measurement campaigns, or to estimate soil moisture otherwise. I miss some critical reflection on this source of error. Reference could be made to the extensively studied cattle-overwintering area in the Czech Republic (Simek and others).

This issue of the representativeness (or lack thereof) of TDR measurements made outside the run, was touched upon in our reply to Referee 1 over his comment on day-to-day variability in fluxes (see above). Referee 2 argues that it would have been feasible to install TDR probes near the sampling points. It certainly would have been desirable from a scientific viewpoint, but their fate and representativeness would have been highly uncertain given the propensity of birds to dig up soil, unearth worms and stones, and other subterranean objects very close to the surface, as attested by the altered and uneven state of the soil surface and the lack vegetation in the close vicinity of chambers. As to the feasibility of setting up TDR probes all over the paddock, this would entail either running hundreds of meters of cables to the different probes, or setting up independent stations near each chamber, which would also inevitably attract the bird's attention, digging, scratching, pecking, defecating and so on.

As suggested in our reply to Referee 1, we have actually addressed in the paper the source of error due to the measurement of WFPS outside the paddock, on p11542 (discussion on gap-filling and effect of WFPS on N₂O emission and on soil respiration) as well as on p11550, l26 onwards.

p. 11535, l. 20 – Do you mean “top left corner of the map”?

Yes, this will be corrected.

p. 11536, l. 1 – Please change to “both interpolation methods”

Yes, this will be changed.

p. 11539, l. 18 – Why are CO₂ fluxes reported in molar, not mass units? It would facilitate a comparison with literature on other grazed systems, which is relevant as these are drawn upon in the modeling of NEE.

The literature on CO₂ fluxes over grasslands, forests, crops, semi-natural vegetation, etc, frequently reports instantaneous (or often half-hourly) NEE, Reco, GPP fluxes in $\mu\text{mol}/\text{m}^2/\text{s}$, but very seldom as $\mu\text{g}/\text{m}^2/\text{s}$, while annual fluxes are expressed as kg or tonnes CO₂-C /ha/yr. Our unit system is thus compatible and comparable with other publications on CO₂.

p. 11539, l. 28 – “is” should read “was”

Yes, this will be corrected.

p. 11541, l. 2 – I believe a discussion of sources of N₂O would be in place (in the discussion section), and it does not appear suitable to deal with the important topic by reference to a general (older) review of processes behind N₂O emissions. A poultry paddock such as this is a highly specialized system with respect to distribution of C and N sources, vegetation cover and, probably, strong gradients in pH with higher pH in the most impacted areas. Both nitrification, denitrification (and nitrifier-denitrification) may be sensitive to pH, and the response of both processes to this particular environment is really difficult to predict, and should be investigated further. At this time I should like to see a short paragraph highlighting the potential importance of C and N concentrating at the surface, and interactions with pH.

We agree that pH is one parameter that one might have expected to vary temporally and spatially within the field as C and N were distributed by the animals. This parameter was indeed measured across the paddock at the start and end of the experiment. The mean soil pH values for the 0-5 cm and 5-15 cm layers (average of 25 composite samples taken across the paddock) were 8.1 and 8.2 at the start of the experiment in 2009, respectively, and they decreased only very

slightly to 7.9 and 8.1 at the end of the experiment in 2010, despite the large quantities of organic material and N deposited. Accounting for uncertainties in sampling and in pH determination, this temporal evolution is not significant in terms of impact on microbial communities involved in nitrification and denitrification. Nor was there any organised spatial variability at the end of the period. In fact the soil at this site is very stable and highly buffered with a high calcium carbonate content (20%), which explains the pH stability and uniformity.

We will add the following short paragraph to Section 4.2, “Spatial and temporal controls of GHG fluxes”, on p11550, l5:

“N₂O emission is known to respond to soil pH, but the effect is complex. The rates of both nitrification and denitrification increase with pH from acidic conditions (pH 3 to 5) to slightly alkaline soils. However, for N₂O emissions conflicting results have been reported, with a positive impact of H⁺ on N₂O emissions in acidic soils, if denitrification is the main source, and a negative impact of H⁺ on emissions in near-neutral to slightly alkaline conditions, if nitrification is the main source (Granli and Bøckman, 1994). The slightly alkaline soil pH at this site (8.1) might suggest a dominant role of nitrification. On the other hand, the change in pH from start to end of this experiment was only of the order of -0.1 pH units, and no significant spatial variability in pH was observed in response to the differential loads of organic matter and N across the paddock, due to the strong stability and buffer provided by the high calcium carbonate content (20%).”

p. 11546, l. 20 – The authors calculate N₂O emission factors based on N excreted, and without correcting for NH₃ losses which are stated to range potentially up to 60% of N excreted. This is in accordance with the recommendations of IPCC for the Tier 1 approach. However, the authors also used a “mechanistically grounded gap-filling algorithm” based on soil temperature and moisture (p. 11541, l. 8) to model emissions between measurement campaigns, which is moving into a higher Tier. The potential for NH₃ losses is also, like N₂O, a function of temperature and moisture at the soil surface. It would be useful to briefly discuss the implications of disregarding NH₃ losses.

We chose Tier 1 approach since we didn’t measure field NH₃ losses due to N excretion on the outdoor run. In our study, it is certainly more relevant to stick to a Tier 1-type approach rather than moving to Tier 2 with an assumed EF for NH₃ losses, which would probably be not representative of our system.

The Tier 1 approach (expression in % of total outdoor N excretion) allows an easy comparison with N₂O emissions in the broiler house, which represented less than 1.5% of indoor N excretion (Meda, 2011), in the perspective of an overall assessment of N₂O emissions at the scale of the whole production system.

To dispel any ambiguity about the context, calculation and applicability of the EF, we have added the following text to Section 4.1 on p 11546, after this sentence:

“For N₂O, the calculation of an EF requires ... control chambers (Table 2). Note that in this paper, the EF we calculate is based on the total N input (as in Tier 1 methodology of IPCC), not on a “net” N input as defined by total input minus NH₃ volatilisation (as in Tier 2 methodology of IPCC), which would obviously result in much higher EF values (~factor 2, if one considers a 50% NH₃ volatilisation/abatement rate).”

p. 11547, l. 4 – In the discussion of CH₄ oxidation it appears relevant to make a reference to the fact that mineral N is often shown to inhibit methane oxidizing bacteria.

This point was made in the paper, on p11550, l1-4.

p. 11547, l. 24 – “more than”?

Yes, we will correct to “...more than 15000...”

p. 11548, l. 18 – “a” should read “an”

Yes, we will change to “..an enhanced risk...”

p. 11549, l. 16 – The argument that microbial processes behind N₂O and CO₂ emissions were probably different because the former varied 100-fold and CO₂ only five-fold is not valid. It could be true if N₂O was predominantly a product of ammonia oxidation, although this is not very likely. The N₂O:N₂ product ratio of denitrification varies dramatically depending on oxygen availability, pH and distance to the surface, and the difference observed could very well be a result of such mechanisms. I suggest that you either avoid this discussion or address it in more detail.

We agree that the argument as it was presented was incomplete and inconclusive. Since the paper needs to be shortened we propose to delete the small section (p11549, l14-17) pertaining to this discussion, as recommended by Referee 2.

p. 11551, l. 12 – FN₂O was defined as ng N₂O/m²/d on p. 11541, l. 18. It is not consistent with the definition given here.

We will add the subscript “a” to change to F_{N₂O,a} and to F_{CH₄,a} in the equation, in the text and in Table 6, with “a” standing for “annual”.

p. 11551, l. 21ff – I must admit that I do not find the extensive calculations to estimate NEE for the outdoor area convincing. There are many assumptions made, for example that C deposition during the batch production cycles can be considered as C sequestration, and C leaching data obtained from very different grazing systems. In the present study, due to gradients in elevation and excretal depositions, there were clearly areas with a high potential leaching. Also, there is no description of how the ranges in daily soil and vegetation intake (Table 5) were taken into account in the modeling. In my view, the calculation of a net annual GHG budget of the outdoor run does not strengthen the paper, which is already too long.

We have addressed a similar comment by Referee 1 (see above), and as stated before we propose to remove (most of) this section from the main body of the paper to an online Supplementary Material document, which will help reduce the paper length. We will still refer to these results, but as part of a short section 4.3 renamed “4.3 Towards a net GHG budget of the outdoor run”, in which we describe the tentative GHG budgets and discuss the most important missing data required for a robust estimate.

p. 11555, l. 8 - The authors point to EC measurements as the most pressing research need. It is difficult too see how such measurements can be adopted in this system where significant spatial heterogeneity in soil conditions and associated emissions is directly linked to the presence of the housing facility. It should be discussed how such measurements could in fact be adopted. I see a much greater scope for improvement of flux estimates by improving the model of emissions by extending the approach adopted here, but including information on pH and actual soil moisture conditions near the soil surface.

We agree that this site would have been a challenging environment for EC measurements, mostly due to size considerations (the paddock was only 50*50m, but there are much larger runs in commercial farms). There is however no doubt that only continuous flux data from an EC system would be able provide the kind of information (namely, annual NEP) that was identified as the weak part by both Referees in our net GHG budget. Dynamic chambers for NEE on this kind of system are out of the question, for reasons already outlined re. the question of soil moisture content measurements. Filtering out the contributions from the housing system from the tower flux measured on the paddock could be achieved by a careful screening of wind directions aided by the use of footprint modelling. A suitable field and

house setup, in which the house would stand at the downwind end of the run (from a dominant wind direction, eg south westerly in our conditions), would have to be found. We did not mean to imply that EC measurements should and could be applied in all situations, but that the choice of study site should take into account the potential for EC flux measurements. Improved modelling is certainly the objective, and as discussed above would certainly benefit from a higher spatial resolution of soil moisture, pH and perhaps soil temperature profiles, but with the logistical and technical constraints previously discussed.