Research Paper

Evaluation of sampling strategies for estimating ammonia emission factors for pig fattening facilities

Tim Ulen\(\text{a,}\)b, Matthijs R.J. Daelman\(\text{c}\), Julio Mosquera\(\text{d}\), Sam Millet\(\text{b}\), Mark C.M. van Loosdrech\(\text{c}\), Eveline I.P. Volcke\(\text{e}\), Herman Van Langenhove\(\text{f}\), Peter Demeyer\(\text{a,}\)∗

\(\text{a}\) Institute for Agricultural and Fisheries Research (ILVO), Technology and Food Science Unit, Burgemeester Van Gansberghelaan 115-1, 9820 Merelbeke, Belgium
\(\text{b}\) Institute for Agricultural and Fisheries Research (ILVO), Animal Sciences Unit, Scheldeweg 68, 9090 Melle, Belgium
\(\text{c}\) Department of Biotechnology, Delft University of Technology, Julianalaan 67, 2628BC Delft, Netherlands
\(\text{d}\) Wageningen UR Livestock Research, PO Box 338, 6700AH Wageningen, Netherlands
\(\text{e}\) Department of Biosystems Engineering, Ghent University, Coupure Links 653, 9000 Gent, Belgium
\(\text{f}\) Department of Sustainable Organic Chemistry and Technology, Ghent University, Coupure Links 653, 9000 Gent, Belgium

**Article info**

**Article history:**
Received 18 May 2015
Received in revised form 14 September 2015
Accepted 23 September 2015
Published online xxx

**Keywords:** Ammonia, Emission factor, Sampling strategy, Animal housing, Monitoring, Pig husbandry

Determining ammonia emission factors (EF) for fattening pig facilities is important from both a regulatory and a research point of view. However, measurements to determine an EF can be time consuming and costly. Several reduced sampling strategies were developed in the past to reduce the costs and measuring time, by taking into account parameters that influence \(\text{NH}_3\) emissions. A methodology to evaluate the precision and accuracy of estimated EFs solely as a function of the sampling frequency and strategy is demonstrated. This evaluation was done by using two long-term, high frequency datasets which both contained measurements during two consecutive pig fattening periods. These datasets were subjected to simulated sampling strategies. Long-term, low-frequency grab sampling proved to be more accurate than short-term monitoring. Repetitive short-term sampling events result in increased precision, but as this entails higher investment in time and money it is imperative to strike the balance between desired precision and available resources. A method to help as set guidelines to decide upon the number of short-term sampling events or the length of a long-term, low-frequency monitoring strategy is presented.

© 2015 IAgE. Published by Elsevier Ltd. All rights reserved.
1. Introduction

Animal husbandry has an adverse impact on the environment with ammonia (NH₃) as one of the major pollutants. Emissions of NH₃ to the atmosphere, and deposition in the environment, can cause acidification and eutrophication (Cole, Todd, & Wing, 2000; Fangmeier, Hadwigerfangmeier, Vandereerden, & Jager, 1994). Therefore, NH₃ was the first gas in agriculture subjected to mandatory emission reductions. As a consequence, low-ammonia-emission (LAE) housing systems were introduced in Flanders since 2004. Pig and poultry farmers in Flanders are obliged to use officially approved LAE housing systems when renovating, expanding or building new animal housing. Innovative farmers can also ask permission to build new LAE housing systems but the reduction potential towards NH₃ has not yet been established through measurements. The NH₃ emission from these housing systems has to be measured by an officially approved (research) institute in order to determine an emission factor (EF). This is traditionally done by frequently measuring (at least every hour) NH₃ concentrations and ventilation rates over long periods (>200 d for fattening pigs), covering both warm and cold seasons. However, the costs associated with this methodology are high because of the expensive equipment required and the man-hours involved (up to € 50,000) (Dekock, Vranken, Gallmann, Hartung, & Berckmans, 2009).

To reduce these high costs, several researchers have tried to reduced sampling strategies for the determination EFs for NH₃. In the Netherlands, a slightly reduced sampling protocol was developed under the Green Label framework (Groen Label, 1996). The goal of this sampling protocol was to accurately estimate the mean annual NH₃ emission of a housing system. This protocol still had a very elaborate sampling protocol with measurements over two growth cycles (e.g. fattening periods), one in summer and one in winter, both on the same farm. Ammonia concentrations had to be measured continuously (i.e. every 5–10 min). Afterwards, hourly means were used in further calculations (Groen Label, 1996). This protocol is also currently used in Flanders to determine NH₃ emission factors. This extensive, more expensive and time consuming protocol was followed-up in The Netherlands using an alternative sampling protocol (a multiple-location approach), based on measurements at several (i.e. four) farms provided with the same housing system. This new protocol prescribes six sampling periods of 24 h for each farm location, distributed over the year and randomly taken over a period of two consecutive month period. For animal categories with growth production cycles (e.g. fattening pigs) measurements had to be equally divided over the production cycle (e.g. fattening period) (Ogink, Mosquera, & Hol, 2011). It was estimated that the total measurement error (expressed as standard deviation) for this new sampling protocol ranged between 15 and 20% (Mosquera, Hol, & Ogink, 2008; Ogink, Mosquera, & Melse, 2008). Recently, an alternative sampling protocol (a case-control approach) has been suggested (Ogink, Mosquera, & Hol, 2013). This approach is based on performing simultaneous measurements in both a newly proposed housing system (or in an existing housing system, but with application of a new management strategy; referred as “case” in this protocol) and a reference system (with known emission factor), both located at the same farm. The number of measurement periods per farm (six) and the conditions concerning spreading of the measurements over time remained the same, but the number of farms decreased from four to two. On the basis of the difference between the emissions from the reference system and the new housing system, the emission factor of the new housing system is estimated. These two alternative protocols (multiple-location approach and case-control approach) are incorporated in the international VERA protocols (VERA, 2011). Mosquera and Ogink (2011) investigated two alternative approaches to shorten the sampling protocol within the four farms. The sampling period was shortened from one year to six months, resulting in only three 24-h sampling periods per farm. Using the results from those six months, either solely or in combination with a mathematical model, led to a small increase in overall random measurement error of the mean emission (between 15.4 and 20.3% with the mathematical model, between 16.8 and 21.4% without the mathematical model) (Mosquera & Ogink, 2011).

Along with the protocols to determine housing system-specific emission factors, reduced building-specific NH₃ sampling strategies for fattening pigs were proposed by Vranken, Claes, Hendriks, Darius, and Berckmans (2004) and later refined by Dekock et al. (2009). In the final protocol, a linear model containing ventilation rate, mean weight of the animals and inside and outside temperature, measured at specific times, was used to model the NH₃ emission from a building. In total, four measurement periods (2 before day 70 and 2 after day 70 in the fattening period) per fattening period were needed to get a good estimate of the NH₃ emissions (i.e. a maximum deviation of less than 10% between the measured and simulated NH₃ emission). To get an EF for the building, three fattening periods had to be monitored (Dekock et al., 2009; Vranken et al., 2004).

Up to now, all reduced sampling strategies suggested in the literature take into account the parameters that influence NH₃ emissions, such as the increasing live weight of the pigs during a fattening period and the seasonal variations in NH₃ emissions. A different method, that does not take into account influencing parameters, was recently published for

---

**Nomenclature**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Number of animal places</td>
</tr>
<tr>
<td>Ci</td>
<td>Incoming NH₃ concentration (mg m⁻³)</td>
</tr>
<tr>
<td>Co</td>
<td>Outgoing NH₃ concentration (mg m⁻³)</td>
</tr>
<tr>
<td>ε</td>
<td>Relative error</td>
</tr>
<tr>
<td>ER</td>
<td>Emission rate (g h⁻¹)</td>
</tr>
<tr>
<td>EF</td>
<td>Emission Factor (kg NH₃ year⁻¹ (animal place)⁻¹)</td>
</tr>
<tr>
<td>LAE</td>
<td>Low-Ammonia-Emission</td>
</tr>
<tr>
<td>μ</td>
<td>Average</td>
</tr>
<tr>
<td>N</td>
<td>Number of emission rates</td>
</tr>
<tr>
<td>n</td>
<td>Number of sampling instances</td>
</tr>
<tr>
<td>Q</td>
<td>Ventilation rate (m³ h⁻¹)</td>
</tr>
<tr>
<td>σ</td>
<td>Standard deviation</td>
</tr>
</tbody>
</table>

**1. Introduction**

Animal husbandry has an adverse impact on the environment with ammonia (NH₃) as one of the major pollutants. Emissions of NH₃ to the atmosphere, and deposition in the environment, can cause acidification and eutrophication (Cole, Todd, & Wing, 2000; Fangmeier, Hadwigerfangmeier, Vandereerden, & Jager, 1994). Therefore, NH₃ was the first gas in agriculture subjected to mandatory emission reductions. As a consequence, low-ammonia-emission (LAE) housing systems were introduced in Flanders since 2004. Pig and poultry farmers in Flanders are obliged to use officially approved LAE housing systems when renovating, expanding or building new animal housing. Innovative farmers can also ask permission to build new LAE housing systems but the reduction potential towards NH₃ has not yet been established through measurements. The NH₃ emission from these housing systems has to be measured by an officially approved (research) institute in order to determine an emission factor (EF). This is traditionally done by frequently measuring (at least every hour) NH₃ concentrations and ventilation rates over long periods (>200 d for fattening pigs), covering both warm and cold seasons. However, the costs associated with this methodology are high because of the expensive equipment required and the man-hours involved (up to € 50,000) (Dekock, Vranken, Gallmann, Hartung, & Berckmans, 2009).

To reduce these high costs, several researchers have tried to reduced sampling strategies for the determination EFs for NH₃. In the Netherlands, a slightly reduced sampling protocol was developed under the Green Label framework (Groen Label, 1996). The goal of this sampling protocol was to accurately estimate the mean annual NH₃ emission of a housing system. This protocol still had a very elaborate sampling protocol with measurements over two growth cycles (e.g. fattening periods), one in summer and one in winter, both on the same farm. Ammonia concentrations had to be measured continuously (i.e. every 5–10 min). Afterwards, hourly means were used in further calculations (Groen Label, 1996). This protocol is also currently used in Flanders to determine NH₃ emission factors. This extensive, more expensive and time consuming protocol was followed-up in The Netherlands using an alternative sampling protocol (a multiple-location approach), based on measurements at several (i.e. four) farms provided with the same housing system. This new protocol prescribes six sampling periods of 24 h for each farm location, distributed over the year and randomly taken over a period of two consecutive month period. For animal categories with growth production cycles (e.g. fattening pigs) measurements had to be equally divided over the production cycle (e.g. fattening period) (Ogink, Mosquera, & Hol, 2011). It was estimated that the total measurement error (expressed as standard deviation) for this new sampling protocol ranged between 15 and 20% (Mosquera, Hol, & Ogink, 2008; Ogink, Mosquera, & Melse, 2008). Recently, an alternative sampling protocol (a case-control approach) has been suggested (Ogink, Mosquera, & Hol, 2013). This approach is based on performing simultaneous measurements in both a newly proposed housing system (or in an existing housing system, but with application of a new management strategy; referred as “case” in this protocol) and a reference system (with known emission factor), both located at the same farm. The number of measurement periods per farm (six) and the conditions concerning spreading of the measurements over time remained the same, but the number of farms decreased from four to two. On the basis of the difference between the emissions from the reference system and the new housing system, the emission factor of the new housing system is estimated. These two alternative protocols (multiple-location approach and case-control approach) are incorporated in the international VERA protocols (VERA, 2011). Mosquera and Ogink (2011) investigated two alternative approaches to shorten the sampling protocol within the four farms. The sampling period was shortened from one year to six months, resulting in only three 24-h sampling periods per farm. Using the results from those six months, either solely or in combination with a mathematical model, led to a small increase in overall random measurement error of the mean emission (between 15.4 and 20.3% with the mathematical model, between 16.8 and 21.4% without the mathematical model) (Mosquera & Ogink, 2011).

Along with the protocols to determine housing system-specific emission factors, reduced building-specific NH₃ sampling strategies for fattening pigs were proposed by Vranken, Claes, Hendriks, Darius, and Berckmans (2004) and later refined by Dekock et al. (2009). In the final protocol, a linear model containing ventilation rate, mean weight of the animals and inside and outside temperature, measured at specific times, was used to model the NH₃ emission from a building. In total, four measurement periods (2 before day 70 and 2 after day 70 in the fattening period) per fattening period were needed to get a good estimate of the NH₃ emissions (i.e. a maximum deviation of less than 10% between the measured and simulated NH₃ emission). To get an EF for the building, three fattening periods had to be monitored (Dekock et al., 2009; Vranken et al., 2004).

Up to now, all reduced sampling strategies suggested in the literature take into account the parameters that influence NH₃ emissions, such as the increasing live weight of the pigs during a fattening period and the seasonal variations in NH₃ emissions. A different method, that does not take into account influencing parameters, was recently published for...
2. Materials and methods

2.1. Datasets

Two datasets were used to evaluate the influence of reduced sampling strategies on the estimated NH₃ emission factor from fattening pig facilities. The datasets originated from measurements performed to assign Green Label certificates for different animal housing systems in the Netherlands, applying the official Dutch measurement protocol for ammonia emission factors.

The first dataset was collected in a conventional (i.e. non-LAE) fattening pig facility between March 13th, 1996 and November 18th, 1996 during two consecutive fattening periods (Fig. 1A & B). During the measurements, 130 fattening pigs were present in the pig unit. The dataset consisted of 5628 hourly outgoing (exhaust) and incoming (inlet) NH₃ concentrations (mg m⁻³) and ventilation rates (m³ h⁻¹). The second dataset was collected in a LAE fattening pig facility between January 28th, 2001 and September 18th, 2001 during two consecutive fattening periods (Fig. 1C & D). During the measurements, 144 fattening pigs were present in the pig unit. The dataset consisted of 5338 hourly outgoing NH₃ concentrations (mg m⁻³) and ventilation rates (m³ h⁻¹). No information on the incoming NH₃ concentrations (background level) was present, but it can be reasonably assumed that the ingoing air did not contain ammonia.

In the first dataset, for each hour, the NH₃ emission rate (ER, g h⁻¹) was calculated according to Eq. (1). The average NH₃ EF (kg year⁻¹ animal place⁻¹) was then calculated (Eq. (2)) by taking the mean over all the (N) NH₃ emission rates and converting this mean from g h⁻¹ to kg year⁻¹ (animal place)⁻¹. In the second dataset, the hourly NH₃ emission rates and the average NH₃ EF were calculated in the same way, except that no correction for incoming NH₃ concentrations was made.

\[
ER = Q \times \frac{C_0 - C_i}{1000} \times 10^{-3} \quad (1)
\]

\[
EF = \frac{\sum ER}{N} \times \left( \frac{10^{-3} \times 24 \times 365}{a} \right) \quad (2)
\]

With ER: emission rate (g h⁻¹), Q: ventilation rate (m³ h⁻¹), C₀: outgoing NH₃ concentration (mg m⁻³), Cᵢ: incoming NH₃ concentration (mg m⁻³), a: number of animal places and N: number of emission rates.

Based on these calculations, the pig unit in dataset 1 had an EF of 2.16 kg NH₃ year⁻¹ (animal place)⁻¹ and the pig unit in dataset 2 had an EF of 1.01 kg NH₃ year⁻¹ (animal place)⁻¹.

![Fig. 1 - Daily NH₃ emission for both housing systems and both fattening periods.](image-url)
2.2. Evaluation of reduced sampling strategies

Five different reduced frequency sampling strategies were investigated: single grab sampling, continuous sampling for 24 h, continuous sampling for 48 h, 7-d continuous sampling and weekly grab sampling. Each of these sampling strategies was simulated by subjecting the datasets described in Section 2.1 to the five sampling strategies, as in Daelman et al. (2013).

2.2.1. Continuous sampling strategies of 24 h and 48 h

The datasets consisted of hourly emission data. Any average of 24 or 48 consecutive data points of the original dataset is a possible outcome of a 24-h or 48-h continuous sampling strategy. Yet, it was assumed that such a sampling strategy is only performed on working days and that it should start (and end) between working hours (9 a.m.–5 p.m.). Considering these restrictions, the first dataset yielded 115 possible estimates for a 24-h strategy and 72 possible estimates for a 48-h strategy. For the second dataset, there were 116 possible estimates of a 24-h strategy and 85 possible estimates for a 48-h strategy. All possible estimates were collected in histograms with a bin width of 0.25 kg NH₃ year⁻¹ (animal place)⁻¹. The dispersion of these histograms was a measure for the precision of the simulated sampling strategy, while the skewness of the histograms indicates its accuracy.

2.2.2. Continuous sampling strategy of 1 week

Similar to the 24-h and 48-h strategy, any average of 168 (7 × 24) consecutive hourly data points of the original dataset is a possible outcome of a 7-d continuous sampling strategy. Again, it was assumed that such a sampling strategy could only start on working days and that it should start (and end) between working hours (9 a.m.–5 p.m.). Considering these restrictions, there were 138 and 142 possible estimates for the two datasets, respectively. All possible estimates were collected in histograms with a bin width of 0.25 kg NH₃ year⁻¹ (animal place)⁻¹.

2.2.3. Single grab sample

For the single grab sample strategy, it was assumed that such a grab sample was only taken during working hours (9 a.m.–5 p.m.) on working days. Therefore, all hourly values that met those restrictions are a possible emission estimate resulting from a single grab sampling strategy. For dataset 1, there were 1244 possible estimates, while there were 1258 possible estimates for dataset 2. All possible estimates were collected in histograms with a bin width of 0.25 kg NH₃ year⁻¹ (animal place)⁻¹.

2.2.4. Long-term weekly grab sampling

In a long-term weekly grab sampling strategy, a single grab sample was taken e.g. every Monday, for a certain number of weeks. In that case, the average of these samples is an estimate of the emission. This was mimicked as follows. For each Monday in the original dataset, a data point between 9 a.m. and 5 p.m. was randomly picked. This was done for each working day. Since not all days of the week are equally represented in the original dataset, not all simulated sampling strategies had the same length: the lengths varied between 31 and 34 weeks. The average of these 31 to 34 (depending on the day of the week) randomly picked data points is a possible emission estimate for a long-term weekly sampling strategy.

Since there are 8 data points available per day and since a sampling strategy lasts for about 30 weeks, the number of possible outcomes is very high (8³⁰). Therefore, the procedure was repeated 1000 times for each week day in order to get a representative sample of the full number of possible estimates. Doing so, 1000 estimated NH₃ EFs were obtained for each week day. This procedure was also applied to the scenario where a weekly sample was taken on a random working day instead of on a fixed day, again repeated 1000 times. For each of the fixed and random weekday strategies, the 1000 values were possible outcomes for the respective sampling strategy. All possible estimates were collected in histograms with a bin width of 0.25 kg NH₃ year⁻¹ (animal place)⁻¹.

2.3. Precision and accuracy as function of sampling frequency

Each of the short-term sampling strategies described above (single grab sample, 24-h and 48-h continuous monitoring and 7-d continuous monitoring) could be repeated, while the long-term weekly grab sampling strategy could vary in length. This would increase the precision of the simulated estimates. The influence of taking multiple (n) random grab samples, 24-h periods, 48-h periods, 7-d periods or weekly grab samples during n consecutive weeks on the estimated EF was evaluated. These evaluations started by calculating all possible grab samples, 24-h periods, 48-h periods, 7-d periods or long-term weekly grab sample strategies that fulfilled the requirements as mentioned in Section 2.2. Thereafter 1 to n random grab samples, 24-h periods, 48-h periods or 7-d periods or long-term weekly grab sample strategies for 1 to n consecutive weeks were randomly taken for all possible periods. Finally, for all 1 to n sampling cases, the relative error between the true EF (as determined with Eq. (2) and based on the whole dataset) and the estimated EF (based on a limited number of measurements and only representing an estimation of the true EF) was determined as a function of the number of sampling cases for each of the investigated reduced frequency sampling strategies (Daelman et al., 2013). These sampling cases were the number (n) of 24-h (or 48-h) random periods for the 24-h (or 48-h) reduced sampling strategies, or the number of random, 7-d periods for the one week reduced sampling strategy. For the random grab sampling strategies, the number of single grab samples was determined necessary to obtain an estimate with a relative error smaller than ±15%. Similarly, for the long-term weekly grab sampling strategy, the number of weeks necessary to obtain an estimate with a relative error less than ±15% was determined.

The relative error (ε) was calculated for all sampling strategies using the following equation:

\[
ε = \frac{\text{Estimated emission factor} - \text{True emission factor}}{\text{True emission factor}}
\]  

(3)

For the 24-h/48-h sampling strategy, this relative error could be calculated for any estimated EF, based on a number (n) of 24 h/48 h periods. These n 24-h/48-h periods were
randomly picked from all possible 24-h/48-h periods calculated in Section 2.2.1 and \( n \) could not be greater than the maximum number of possibilities calculated in Section 2.2.1. These periods did not have to be consecutive. An example of a plot of the relative error \( \varepsilon \) against the number \( (n) \) of 24-h periods is seen in Fig. 2.

Following a similar approach as described above for the 24-h and 48-h sampling periods, the effect of the number of 7-d periods was assessed. As a result, 138 (dataset 1) and 142 (dataset 2) 7-d moving averages were calculated, representing all possible 7 consecutive measuring day periods starting between 10 and 11 a.m. and not starting in weekends. The maximum of non-overlapping periods in dataset 1 and 2 were 27 and 28, respectively. Thus, a monitoring strategy could exist of \( n \) random non-overlapping and not necessarily consecutive weeks, with a maximum of \( n = 27 \) (dataset 1) or \( n = 28 \) (dataset 2). As for the 24-h and 48-h sampling strategy, all of the hourly ERs over the \( n \) 7-d periods were averaged for each \( n \).

A grab sampling strategy could consist of 1 to \( n \) grab samples. However, there were 1244 and 1258 possible single grab samples between 9 a.m. and 5 p.m. on working days in respectively dataset 1 and dataset 2. Taking a lot of grab samples is not very practical. In order to mimic a practical grab sampling strategy, 1 to 300 random grab samples were taken from the 1244 or 1258 possible grab samples. For each \( n \), the \( n \) grab samples were averaged to obtain an estimated ER and then converted to kg year\(^{-1}\) (animal place\(^{-1}\)). This resulted in 300 estimates, one for each value of \( n \).

The long-term weekly grab sampling strategy was simulated by taking weekly grab samples on a random working day. So, this means that a grab sample would be taken on \( n \) consecutive weeks on a random working day. This was simulated by combining the daytime values of the weekdays per week and taking \( n \) weekly grab samples from that combination during 1 to \( n \) weeks, starting with a random week. For each value of \( n \), the values from the \( n \) weekly grab samples were averaged and then converted to kg year\(^{-1}\) (animal place\(^{-1}\)). This resulted in 35 (dataset 1) and 34 (dataset 2) estimated EFs.

For each of the simulated sampling strategies and for each value of \( n \), 1000 simulations were executed, yielding a distribution of the relative errors as a function of the number of sampling periods. Subsequently, the average (\( \mu \)) and the standard deviation (\( \sigma \)) of this distribution were calculated for each value of \( n \). 95% confidence interval around the mean for \( \varepsilon \) are given by \( \mu - 2\sigma \) and \( \mu + 2\sigma \). The confidence intervals of the relative error are a measure of the precision of the simulated sampling strategy, while the average of the relative error represents the accuracy of the estimate.

3. Results

3.1. Effect of the different sampling strategies

Similar outcomes were found for the single grab sampling strategy (Fig. 3), the 24-h sampling strategy (Fig. 4), the 48-h sampling strategy (Fig. 5) and the 7-d sampling strategy (Fig. 6). Except for the single grab sampling strategy in dataset 1, the chance of obtaining an estimated EF that is within 15% of the true EF was higher than the chance of obtaining an estimated EF that is more than 15% higher than the true EF (Table 1). Except for the single grab sampling strategy in dataset 2, these sampling strategies had a higher chance of estimating an EF that is more than 15% higher than the true EF, than estimating an EF that is more than 15% lower than the true EF. For the long-term weekly grab sampling strategy, the chance of obtaining an estimated EF that is within 15% of the true EF was 100% for both datasets and for all days (data not shown).

For the single grab, the 24-h, the 48-h and the 7-d sampling strategy, the precision (defined as the dispersion of the histograms) for the results from dataset 2 was higher than for the results from dataset 1. For the long-term weekly grab sampling strategy, no apparent difference in precision could be found between both datasets.
3.2. Precision as function of sampling frequency

Performing a sampling strategy based on 10 random grab samples, 24-h, 48-h or 7-d sampling periods or weekly grab samples for 10 consecutive weeks yielded an estimate of the EF with a relative error between maximum $-39\%$ and $48\%$ (Table 2). If the number of random 24-h, 48-hr or 7-d sampling periods was increased, the relative error decreased (Figs. 7–9).

The estimate of the EF, based on 20 random 24-h sampling periods, had relative errors in a 95% confidence interval between $-12\%$ and $14.8\%$ (dataset 1) and between $-10\%$ and $18\%$ (dataset 2). Sampling 20 random 7-d periods reduced the relative errors further to $-6\%$ and $7\%$ (dataset 1) and $-4\%$ and $11\%$ (dataset 2).

In order to reduce the relative error to maximum $\pm 15\%$ for the single grab sampling strategy (Fig. 10), 84 (dataset 1) and 27 (dataset 2) random grab samples need to be taken. Taking 100 random grab samples reduced the relative error further down to $-1\%$ and $14\%$ (dataset 1) and $-8\%$ and $7\%$ (dataset 2). If more grab samples were taken, the precision only improved slightly. In order to reduce the relative error to maximum $\pm 15\%$ for the long-term weekly grab sampling strategy (Fig. 11), weekly grab samples had to be taken during 32 (dataset 1) and 28 (dataset 2) consecutive weeks.

For the 24-h and 48-h sampling strategies (Figs. 7 and 8), all 4 graphs showed a small positive bias. This bias is due to the fact that these sampling strategies only take into account hour
Fig. 5 – Histograms of all possible outcomes from the 48-h sampling strategy for dataset 1 (left) and dataset 2 (right). The vertical lines indicate the true NH3 EF.

Fig. 6 – Histograms of all possible outcomes from the 7-d period sampling strategy for dataset 1 (left) and dataset 2 (right). The vertical lines indicate the true NH3 EF.

**Table 1** – Percentage of values lower than 85% of the true EF, between 85% and 115% of the true EF and larger than 115% of the true EF.

<table>
<thead>
<tr>
<th>Sampling protocol</th>
<th>&lt;0.85 × True EF</th>
<th>≥ 0.85 × True EF &amp; ≤ 1.15 × True EF</th>
<th>&gt;1.15 × True EF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single grab</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sampling</td>
<td>dataset 1</td>
<td>26.8</td>
<td>28.5</td>
</tr>
<tr>
<td>dataset 2</td>
<td>32.8</td>
<td>35.6</td>
<td>31.6</td>
</tr>
<tr>
<td>24-h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sampling</td>
<td>dataset 1</td>
<td>25.2</td>
<td>38.3</td>
</tr>
<tr>
<td>dataset 2</td>
<td>25.0</td>
<td>44.0</td>
<td>31.0</td>
</tr>
<tr>
<td>48-h</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sampling</td>
<td>dataset 1</td>
<td>22.2</td>
<td>40.3</td>
</tr>
<tr>
<td>dataset 2</td>
<td>22.4</td>
<td>45.9</td>
<td>31.8</td>
</tr>
<tr>
<td>7-d sampling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>dataset 1</td>
<td>23.2</td>
<td>43.0</td>
</tr>
<tr>
<td></td>
<td>dataset 2</td>
<td>25.4</td>
<td>41.3</td>
</tr>
</tbody>
</table>

**Table 2** – Relative error intervals of an estimated EF, based on 10 random grab samples, 24-h, 48-h or 7-d sampling periods or weekly grab samples for 10 consecutive weeks.

<table>
<thead>
<tr>
<th>Dataset 1</th>
<th>Dataset 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single grab</td>
<td>between –20% and 33%</td>
</tr>
<tr>
<td>24-h</td>
<td>between –21% and 23%</td>
</tr>
<tr>
<td>48-h</td>
<td>between –19% and 21%</td>
</tr>
<tr>
<td>7-d</td>
<td>between –15% and 18%</td>
</tr>
<tr>
<td>Long-term weekly</td>
<td>between –37% and 48%</td>
</tr>
</tbody>
</table>

* The relative error has a 95% chance of lying between these values.
measurements started on specific days, while the true EF is based on continuous data over all days of the week. Furthermore, the lower and upper uncertainty bounds coincided for all 4 graphs. This is because, for each of the 1000 iterations, there is only one way to sample the maximum of n 24-h/48-h sampling periods. When the maximum number of n 24-h/48-h sampling periods was sampled, only one estimate of the EF (and associated relative error) was found. In contrast to the previous sampling strategies, the lower and upper uncertainty bounds did not coincide for both graphs (Fig. 9) when the maximum number of 7-d periods was reached. This is because a 27 or 28-week sampling strategy can start on each working day of the week. Therefore, the estimates between simulated strategies will differ depending on which day of the week the strategy started. If the maximum number of single grab samples would be taken into account, the lower and upper uncertainty bounds would coincide for both graphs (Fig. 10). As for the 7-d sampling strategy, the lower and upper uncertainty bounds for the weekly grab sampling strategy (Fig. 11) did not coincide since the number of possible ways to conduct a weekly grab sampling strategy of 35 (dataset 1) or 34 (dataset 2) consecutive weeks is almost infinite.

4. Discussion

4.1. Evaluation of the different sampling strategies

Estimations of the NH₃ EF based on just one single grab sample, one 24-h period, one 48-h period or one 7-d period
exhibited a relatively high chance (>30%) to overestimate the true EF. At first glance it seems striking that generally, with the exception of the single grab sampling strategy in dataset 2, there was a higher chance to overestimate the true EF than to underestimate it.

When taking a closer look at the initial dataset, multiple explanations for this outcome can be found. It can be seen in Fig. 1 that the NH3 emission increased more drastically towards the end of the fattening periods for both housing systems (although this was not as pronounced in all fattening periods). For example, the NH3 emission in fattening period 1 for the conventional housing system (Fig. 1A) increased rather abruptly between approximately day 30 and day 50. When only one 24-h or 48-h sample is taken during that fattening period, there is a higher chance for that sample to be taken in a period with relatively higher NH3 emission since there are more periods with relatively higher NH3 emission as compared to periods with relatively lower NH3 emission. If an EF was be estimated only on the basis of that sample, the estimated EF has, consequently, a higher chance to be larger than the true EF.

The same explanation applies for the one week sampling strategy. Since there is a higher chance within a one week period of taking in a period with higher NH3 emission, the estimated EF during that period has a higher chance to overestimate the true EF. This explanation is also valid for the single grab sampling strategy. However, for this reduced sampling strategy, the limited time period when the sampling
can be performed (between 9 a.m. and 5 p.m.) also plays a role. During these hours, NH₃ emissions are higher than during the night (Fig. 12). Higher NH₃ emissions during daytime as compared to night time have also been found by other researchers Aarnink, vandenBerg, Keen, Hoeksma, & Verstegen, 1996; Blanes-Vidal, Hansen, Pedersen, & Rom, 2008; Ngwabie, Jeppsson, Nimmermark, & Gustafsson, 2011) and probably originate from increased urination behaviour of the pigs (Aarnink et al., 1996) and higher air movements over the manure surface during daytime (Blanes-Vidal et al., 2008). Since the true EF takes into account both (higher) day and (lower) night values, it is not surprising that, depending on the dataset, approximately 30% or 45% of the grab samples overestimated the true EF with more than 15%.

The absolute difference between day and night values for both systems (Fig. 12) may explain the different results between the two datasets for all short-term sampling strategies. The smaller difference between day and night values for dataset 2 (LAE) gave rise to a better precision (defined as the dispersion of the histograms) as compared to dataset 1 (conventional).

The long-term weekly grab sampling strategy takes a grab sample every week and is therefore better equipped to capture the increase in NH₃ emissions during a fattening period. However, as with the single grab sampling strategy, samples for the long-term weekly grab sampling strategy were only taken during daytime (between 9 a.m. and 5 p.m.). Nevertheless, for all days and both datasets, all of the 1000 simulated

Fig. 11 – Uncertainty bounds of the relative error (ε) as a function of the number (n) of weekly grab samples, based on 1000 iterations.

Fig. 12 – Mean emission rate for days 80–100 in fattening period 2 for the conventional (left) and LAE (right) housing system. Error bars indicate the standard deviations.
long-term weekly grab sampling strategies estimated an EF that is within 15% of the true EF. This good precision for the long-term weekly grab sampling strategy originates from the fact that the increase in NH₃ emissions during a fattening period is well captured. This indicates also that taking grab samples on a regular basis (on a specific day and/or once every week) during daytime can be sufficient to obtain an estimate of the EF that is within 15% of the true value.

After evaluating all sampling strategies, one can decide that the use of only one grab sample, one 24-h sample, one 48-h sample or one 7-d sample has a relatively high chance of deviating from the true value. So, much caution has to be taken when estimating an EF based on only one sample of a short-term sampling strategy. Furthermore, since the spread on the estimated EFs was larger for dataset 1, the use of only one sample is certainly not advisable in situations where large differences in NH₃ emissions between days and between hours exist. Of course, in practice, NH₃ emissions are not determined on such short periods. Therefore, in Section 3.2, the precision and accuracy as a function of the number of samples was determined. The estimated EF based on one long-term weekly grab sample strategy was, in contrast to the other sampling strategies, much more precise and accurate. With one single long-term weekly grab sampling strategy, the estimated EF is within 15% of the true EF.

4.2. Precision as function of sampling frequency

It was shown that the long-term weekly grab sampling strategy yielded the best estimate of the true EF. However, this sampling strategy took the maximum number of weeks into account. Since sampling for such a long period would be costly, simulations were made to determine if a shorter sampling frequency (i.e. taking grab samples for fewer weeks) would also be sufficient. Since weekly grab sampling on a random working day yielded equally good results as weekly grab sampling on a specific day of the week, it was decided to test the shorter sampling frequency for a random working day only. Decreasing the number of consecutive sampling weeks increased the relative error gradually. For 15 weeks of sampling, the relative error was between −25% and 25% (Fig. 11). With less than 15 sampling weeks, the relative error increased rapidly. If grab sampling was performed for 25 consecutive weeks on a random working day, an EF was estimated with a relative error between −12% and 25% for dataset 1 and a relative error which has 95% chance of lying between −17% and 13% for dataset 2. If the same number of grab samples was taken completely at random (i.e. not systematically one measurement every week), the resulting estimated EF had a relative error which has 95% chance of lying between −10% and 24% (dataset 1) and between −16% and 15% (dataset 2). So, a similar precision was obtained with completely random grab samples as with grab samples every week. If the number of random grab samples or the number of consecutive weeks decreased even further, the random grab sampling strategy had a better precision as compared to grab sampling each week (Figs. 10 and 11).

If more information on the diurnal patterns is desirable, strategies based on grab samples cannot be used. Instead, the 24-h, 48-h or 7-d measuring strategy has to be used. These strategies can also be used to estimate an EF if multiple 24 h, 48 h or 7 d periods are taken. To get an estimated EF with a relative error below 15%, 21 (dataset 1) or 27 (dataset 2) 24-h periods, 20 (dataset 1) or 29 (dataset 2) 48-h periods or 13 (dataset 1) or 15 (dataset 2) 7-d periods were necessary. It can be concluded that there is no large difference in the number of 24-h or 48-h periods that have to be sampled. The number of 7-d periods needed to get a good estimate of the EF is about 40% lower than the number of 24-h or 48-h periods.

When deciding which strategy to use, the efforts and time needed to install and dismantle the measuring setup should also be taken into account. Besides the costs benefits for the farmers and companies, reducing the sampling time also ensures a higher availability of the measuring equipment. This allows multiple measuring strategies and assessments of mitigation techniques to be performed during a year.

All of the simulated sampling strategies used in this study were based on only two datasets. Therefore, the results and conclusions drawn in this study should be further tested with other datasets including those obtained from other buildings. This could be verified in a follow-up study where the same sampling strategies are applied to other long-term continuous datasets.

Another very important remark to make is that the sampling strategies proposed here, are only valid to estimate an EF for a certain barn and cannot be generalised to all buildings using the same housing system. In contrast, Ogink et al. (2011) and Mosquera and Ogink (2011) proposed a reduced sampling strategy based on a limited number of measurements at multiple farms, in order to take into account the between-farm variance. Furthermore, these measurements have to be spread out over time.

5. Conclusions

Two long-term datasets of high-frequency ammonia emission measurements, from a conventional and a low emission fattening pig housing facility were used to evaluate different reduced sampling strategies. In contrast to previous work on reduced sampling strategies, this study is based on emission data only, disregarding any other parameters that affect the emission.

Short term sampling strategies such a single grab sample, a 24-h or 48-h high frequency monitoring period or a 7-d high frequency monitoring period provide less accurate and less precise emission estimates than a low frequency strategy consisting of weekly grab samples carried out throughout the production cycle, although the practical restrictions that occur because weekly grab samples can only be taken during the working day can cause a bias leading to an overestimation of emissions.

Repeating the short-term high-frequency sampling events throughout the production cycle could alleviate their lack of precision, while the length of the weekly grab sampling strategy could be reduced without sacrificing too much precision. As the cost and effort of monitoring increase with the number of repeated sampling events, or the length of a weekly sampling strategy, the resources required should be balanced against the required precision. The present contribution
presented a method that can be used as a guideline to negotiate this trade-off.

REFERENCES


