Article (refereed) - postprint

Hellsten, S.; Dragosits, U.; Place, C.J.; Dore, A.J.; Tang, Y.S.; Sutton, M.A..
2018. **Uncertainties and implications of applying aggregated data for spatial modelling of atmospheric ammonia emissions.** *Environmental Pollution*, 240. 412-421. https://doi.org/10.1016/j.envpol.2018.04.132

Uncertainties and implications of applying aggregated data for spatial modelling of atmospheric ammonia emissions

S. Hellsten[a,b,c,*], U. Dragosits[a,b], C. J. Place[b], A.J. Dore[a], Y.S. Tang[a], M. A. Sutton[a]

[a]Centre for Ecology and Hydrology Edinburgh, Bush Estate, Penicuik, Midlothian EH26 0QB, UK

[b]School of GeoSciences, The University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK

[c] IVL Swedish Environmental Research Institute Ltd, P.O. Box 5302, SE-400 14 Gothenburg, Sweden

[*]Address for correspondence: Sofie Hellsten, IVL Swedish Environmental Research Institute Ltd, P.O. Box 5302, SE-400 14 Gothenburg, Sweden, Telephone: +46 (0)31 - 725 62 15, Fax: +46 (0)31 – 725 62 90, E-mail: sofie.hellsten@ivl.se

Co-authors e-mail:

- Ulrike Dragosits: ud@ceh.ac.uk
- Chris J. Place: cjp@staffmail.ed.ac.uk
- Anthony J. Dore: todo@ceh.ac.uk
- Yuk Sim Tang: yst@ceh.ac.uk
- Mark A. Sutton: ms@ceh.ac.uk

Abstract

Ammonia emissions vary greatly at a local scale, and effects (eutrophication, acidification) occur primarily close to sources. Therefore it is important that spatially distributed emission estimates are located as accurately as possible. The main source of ammonia emissions is agriculture, and therefore agricultural survey statistics are the most important input data to an ammonia emission inventory alongside per activity estimates of emission potential. In the UK, agricultural statistics are collected at farm level, but are aggregated to parish level, NUTS-3 level or regular grid resolution for distribution to users. In this study, the Modifiable Areal Unit Problem (MAUP), associated with such amalgamation, is investigated in the context of assessing the spatial distribution of ammonia sources for emission inventories.

England was used as a test area to study the effects of the MAUP. Agricultural survey data at farm level (point data) were obtained under license and amalgamated to different areal units or zones:

regular 1-km, 5-km, 10-km grids and parish level, before they were imported into the emission model. The results of using the survey data at different levels of amalgamation were assessed to estimate the effects of the MAUP on the spatial inventory.

The analysis showed that the size and shape of aggregation zones applied to the farm-level agricultural statistics strongly affect the location of the emissions estimated by the model. If the zones are too small, this may result in false emission "hot spots", i.e., artificially high emission values that are in reality not confined to the zone to which they are allocated. Conversely, if the zones are too large, detail may be lost and emissions smoothed out, which may give a false impression of the spatial patterns and magnitude of emissions in those zones. The results of the study indicate that the MAUP has a significant effect on the location and local magnitude of emissions in spatial inventories where amalgamated, zonal data are used.

Capsule: The aggregation level, i.e. the size and shape of the aggregation zones of point data, has a significant effect on the location and local magnitude of emissions in spatial inventories where aggregated point data are used.

Keywords: Modifiable Areal Unit Problem, zone design, aggregated data, ammonia, agricultural survey data, spatial inventories.

1. Introduction

It is commonly accepted that the main sources of uncertainty in spatial pollution emission inventories are in the way models represent reality, and the input data to such models. Sources of uncertainty in non-spatial emission inventories may be in the activity statistics (representing the polluting activity) or the emission potentials (the emission estimated per unit of polluting activity, often referred to as "emission factors"). For many emission inventories, uncertainties in emission potentials and activity data have been estimated by identifying upper and lower limits of certainty (Beusen *et al.*, 2008; Kühlwein and Friedrich, 2000; Misselbrook *et al*., 2000; Sutton *et al*., 1995, 2013; Winiwarter and Rypdal, 2001;

Zheng *et al*., 2012). When emissions are spatially distributed, a further dimension of uncertainty is added, due to the introduction of the spatial dimension to emissions. While uncertainties of the magnitude of emissions have generally been fairly well investigated, uncertainties due to spatial issues tend to have been overlooked in the past, with only a few studies having investigated spatial uncertainties to some extent (Dragosits *et al*., 2002; Leopold *et al.*, 2012; Lindley *et al*., 2000; Winiwarter *et al*., 2003; Xu *et al*., 2016).

Ammonia emissions vary greatly at a local scale and some effects (eutrophication, acidification) occur primarily close to sources. Thus it is important to minimize uncertainties in the spatial location of the estimated ammonia emissions, due to the high spatial variability in atmospheric concentrations and dry deposition of NH_3 (Cellier *et al*., 2011; Dragosits *et al*., 2002 and 2006; Hallsworth *et al*., 2010; Sutton *et al.*, 1998). Errors and uncertainties in the emission map will inevitably have implications on the result of models that use spatial inventories as their main input data, e.g., atmospheric transport and deposition models and assessments of critical loads and critical level exceedance.

1.1. MAUP and Agricultural Survey Data

The main source of ammonia emissions is agriculture, and agricultural survey statistics are the most important input data to a spatial ammonia emission inventory. In the UK, agricultural statistics are collected at a very fine resolution (farm/agricultural holding level), but aggregated to a much coarser resolution e.g. 5-km grid cells, NUTS-3 level (nomenclature of territorial units for statistics (EU, 2003), i.e. counties, unitary authorities, council areas or districts), civil parishes or parish groups, for distribution to users, to ensure individual holdings cannot be identified.

Such anonymity is a legal requirement for these data, which are collected on the basis that data providers will be in no way prejudiced by reporting data. In the past, these spatial resolutions have generally been accepted to provide a reasonable balance between spatial uncertainty and resolution in models (Asman *et al.*, 1998). However, further to the importance of 'hot-spots' for ammonia (e.g. Loubet *et al*., 2009), and that national assessments at the 5 km grid level underestimate the occurrence of critical loads exceedances due to ammonia in agricultural landscapes (Dragosits *et al*., 2002), there has been increasing concern about limits in the spatial resolution applied in ammonia emission inventories. Geels *et al*. (2012) for instance, showed that an increase of resolution improves model results for air pollution transport models. There is, however, little knowledge of the actual effect of the zonal aggregation on the result. When agricultural survey data are aggregated from farm-level (point data) to a coarser resolution (area data), the data are generalised and variability between farms within each zone is lost. In addition, this loss of information is not necessarily consistent from one zone to the other (Openshaw and Rao, 1995). Aggregated data give different results depending on the scale, size, shape and location of the aggregation zones (Dark and Bram, 2007; Openshaw, 1984). This problem is referred to as the Modifiable Areal Unit Problem (MAUP).

Although some spatial emission inventories discuss the problem of the MAUP, e.g. Maes *et al*. (2009), few studies on the effects of the MAUP with regard to emission inventories can be found in literature (e.g., Dai and Rocke, 2000; Lindley *et al.*, 2000). The present study therefore appears to be one of the first research efforts demonstrating effects of the MAUP in the context of spatial emission inventories.

Aggregating the agricultural holding data into zones ensures that information on individual holdings in the survey results will not be identifiable, as required by agencies collecting the data. Geddes *et al*. (2003) suggest that geographical variation in the physical characteristics of the farms and the parishes is the most significant problem in spatial modelling of these types of data. Point data (such as farm holdings) can be difficult to analyse, but when the data have been aggregated into zones, spatial analysis of the data becomes possible. Other advantages of aggregating the data are that geographical patterns are created, and the volume of the data is reduced (Openshaw and Alvadines, 1999). The main disadvantage is that information and spatial detail is lost in the aggregation process.

The term 'modifiable' refers to the fact that the spatial units (the zones) can be changed, and a different distribution would be generated if a different zoning system was used. Aggregation of the data can be achieved in many different ways, both in terms of scale and zone characteristics (Openshaw, 1977). The MAUP is hence mainly associated with two effects (Openshaw, 1984; Openshaw and Taylor, 1979):

- *The scale effect* – the same data may give different results for zones of different sizes.
- *The zonation effect* - results may vary even with the same scale, depending on the location of the zonal boundaries and how the units are aggregated.

1.2 Modelling ammonia emissions

The general methodology to model ammonia emissions is to multiply an emission potential with spatially distributed activity data, such as the agricultural survey statistics. In this study ammonia emissions were modelled at a 1-km grid resolution with the UK AENEID model (Dragosits *et al.*, 1998, Hellsten *et al.*, 2007, 2008). The agricultural survey data for England are normally available at parish level or 5 km x 5 km grid resolution, to avoid identification of individual farms. To calculate a gridded ammonia emissions inventory from irregularly shaped and sized polygons in the UK, landcover data are used as a proxy, to spatially locate emissions within each zone, i.e., by using 'intelligent area weighted interpolation' (Sadahiro, 2000). Introducing a geographical property such as land cover within the parish zones is a means to reduce the spatial representation error within each zone, because ammonia emissions from different agricultural sources tend to occur on specific land cover types. Land cover correlates well with most agricultural data (except non-land-based enterprises such as large intensive pig and poultry farms).

While it is technically easy to aggregate small units into larger units (up-scaling), down-scaling is not possible without additional information (Montello, 2001) or introducing additional uncertainty through expert judgement. When the agricultural survey data for each zone (parish or 5-km grid cell) are re-distributed at a 1-km grid resolution, a spatial representation error is introduced. The magnitude of the error depends on the type of emission (point or area source etc.) as well as the zone size and the location of zonal boundaries (Longley and Batty, 1996).

2. Methodology

In this study, disclosive farm holding data for England were obtained and analysed in relation to the MAUP. This analysis raised issues of data confidentiality, as more could be seen in the disclosive outcomes than is possible to visualize when complying with the requirements of data confidentiality. Further details of the handling of confidentiality of the agricultural datasets in the emission calculations are provided in Hallsworth *et al.* (2010). In addition, all figures representing actual holding data in the current study have been modified, and include up to 10 % additional random data points, thereby ensuring that the output is non-disclosive.

The MAUP and its effects are thus investigated by aggregation of holding data (point data) for England using different zonal systems (Figure 1). Four different zoning systems are tested here. Three gridded systems (1-km, 5-km and 10-km level) were chosen because a regular square pattern facilitates further use and analysis of the data. The fourth zoning system uses irregular polygons, in this case civil parishes, a common aggregation format available to users. Each point (holding) was assigned to the zone in which it is located, for all four zoning systems. The associated agricultural survey data for all holdings in each zone (e.g., livestock populations, crop areas) were aggregated, and each emission source category was re-distributed within each zone at a 1 km x 1 km grid, according to the AENEID methodology (Dragosits *et al.*, 1998, Hellsten *et al.*, 2008). The spatial model allocates emissions from each survey item onto relevant land cover types within the aggregation zone. For instance, housing and storage emissions are assumed to occur on improved pasture (cattle and sheep), suburban areas (poultry) and arable and suburban areas (pigs). Grazing emissions occur on different types of grassland and spreading emissions on arable land and grassland. Ammonia emissions from crops and grassland were distributed onto the landcover types arable land and improved grassland, respectively. If relevant land cover types are not present in an aggregation zone, the model reassigns the survey data to the next most likely land cover type. This is to ensure that no survey items are lost in the modelling process, due to discrepancies between the survey data and the land cover map.

Figure 1

Ammonia emission maps for each zone system were calculated at a 1 km x 1 km grid by applying emission potentials to the distributed source activity output from the AENEID model. The same original activity data and emission potentials were applied for all four scenarios, the only difference being the type of zoning system applied. It follows that although all four scenarios are based on different zonal systems, the magnitude of the total emission will be the same, and any difference in the spatial location of the emissions will be indicative of the effect of the MAUP.

The potential areal extent of ammonia emissions from a single holding allocated to different zoning systems is demonstrated in Figure 2. The areal extent and the spatial location of the emission depend on the size of the zones (the *scale effect*), as well as the location of the zones (the zonal effect). The scale effect can be seen in Figure 2 when comparing the areal extent of the three grids (1-km, 5-km & 10-km). Although the areal extent of the sample parish and the 5-km grid cell containing the sample farm in Figure 2 are almost of the same size, the location of the farm emission is still different because of the location of the zones, demonstrating the *zonation effect*.

Figure2

The gridded emission maps for the four zonation systems were analysed and compared by i) visual interpretation, ii) identifying extreme emission values, and iii) frequency distributions. Furthermore, overestimation of agricultural land within each grid cell due to discrepancies between the survey data and the land cover map was also analysed.

The emission estimates for each zonal system were further evaluated with independent measurements, by comparing modelled NH_3 concentration fields for ammonia, with monitored NH_3 air concentrations from the UK National Ammonia Monitoring Network, NAMN (Sutton *et al.*, 2001a; Tang *et al.*, 2018; Tang and Sutton, 2004). The atmospheric transport model FRAME (Fine Resolution AMmonia Exchange, version 9.15) incorporates the main atmospheric processes (emission, diffusion, chemistry and deposition) to calculate NH_3 concentration fields in the UK (Dore *et al.*, 2007, 2012; Matejko *et al.*, 2009; Vieno, 2006;). FRAME has been found to give a good comparison with measurements of NH_3 concentrations when compared with the performance of other atmospheric chemistry transport models (Dore *et al.*, 2015).

The FRAME model uses maps of ammonia emissions from AENEID as an input, with a livestock sector dependent emission height. Vertical concentration profiles are calculated by simulation of diffusion through 33 layers of varying depth. Loss processes for ammonia represented in the model include chemical conversion to ammonium aerosol, washout by precipitation and dry deposition to surface vegetation.

In the NAMN (http://www.pollutantdeposition.ceh.ac.uk/content/ammonia-network), monthly sampling is carried out at 85 sites using an active diffusion denuder method, the DEnuder for Long-Term Atmospheric sampling system, DELTA (Sutton *et al.*, 2001) and passive diffusion samplers (Tang *et al.,* 2001), with the latter normalized to the former to take account of the effect of the quasi-laminar sublayer of air at the inlet of the passive samplers and other differences in the passive sampler performance. For example Tang *et al*. (2018) have shown that different membrane characteristics at the inlet of passive samplers affect their sampling rate.

3. Results and discussion

3.1 Visual interpretation

Emission maps derived from different aggregation zones at a 1 km x 1 km grid resolution for an example area (400 km^2) are shown in Figure 3. For the 1-km grid zones, the emissions are very scattered and the emission map is characterised by high emission peaks, as well as many grid squares of zero emission. Even though the overall total emission is the same for all four zonal systems, the 1-km zone map gives the impression of lower total ammonia emissions than the other maps, due to the emissions being confined to zones where the holdings are located as points, rather than a realistic distribution of agricultural activities. This is an important issue with agricultural survey data that, although data are collected at holding level (point source), the actual livestock and the agricultural land in reality represent an area source. From observation, most farms, especially larger holdings, or holdings in upland areas, are unlikely to have all their emissions located in the 1 km^2 grid cell where the farm is registered.

Both the 5-km level and the 10-km zone emission maps show a much smoother emission pattern than the 1-km zone map. Many studies of the MAUP have shown that larger aggregation zones tend to have a "smoothing effect" on the result, while smaller zones augment differences. Jelinski and Wu (1996) showed that information on spatial heterogeneity is lost or distorted in the aggregation process, and that the loss of detail increases as the zone size increases. Larger aggregation zones, however, have the advantage of giving more statistically stable distributions, with less extreme and unrealistic emission values. Jelinski and Wu (1996) also showed that, although the size of the zones may remain the same, the variance changes even when the orientation of the zones is changed (the zoning problem).

In Figure 3, the underlying aggregation zones were visible even for the 10-km zone system. In this map a combination of 1 km and 10 km features can be seen as the 1 km resolution redistribution of the AENEID model related to land cover is imposed on the 10 km resolution distribution of agricultural statistics used. The emission map based on a combination of 1 km land cover redistribution and the parish distribution of agricultural statistics did not give rise to any visible artificial borders, which may give the impression that the underlying distribution zones are undetectable. Although the parish borders are in many cases as artificial constructs as the square zones, they are not as easily detectable by eye as square patterns, due to their irregular sizes and shapes. When superimposing the parish boundaries onto the parish zone map, some of the parishes could be identified in the underlying emission map. In summary, this effect varies with location and size of the aggregation zones. For instance, pig and poultry emissions are often located in small parishes in lowland areas, while cattle and sheep emissions tend to be located in larger parishes in upland areas, and hence the underlying aggregation zones are more difficult to detect.

Figure 3

One way to reduce the uncertainties in the spatial distribution even further, could be to deal with large point sources such as pig and poultry production facilities in a particular manner. For instance, these facilities can be identified from satellite images or from the IPPC-database (if access is improved). In the disaggregation process, remote sensing data and particular distribution processes could be applied to narrow down the uncertainties as to where the bulk of the emissions originates. For instance, Hellsten (2006) has suggested an iterative process based on manure saturation rates to distribute pig and poultry manure, independent of the agricultural land within the aggregation zone.

3.2 Extreme emission values

In Table 1, extreme emission values for each of the four zone systems were identified. As expected, the 1-km zone system has the most extreme values. This is followed by the parish zones, the 5-km zones, and finally the 10-km zones. Table 1 clearly shows that emission values are less extreme the larger the aggregation zones (the smoothing effect).

Table1

Rather extreme values were common in the ammonia emission map based on the parish distribution, with the parish zones varying in size between 1 km^2 and 258 km^2 across England. Extreme values tend to occur in small parishes (2 % of all parishes are approximately 1 km^2), but high emission values may also be present in large parishes, if the relevant land cover associated with an emission source type in that parish is limited to just a few grid cells.

The emission map based on 1-km aggregation zones clearly showed the problems of hot spots as well as grid squares lacking any ammonia emissions. Similarly, parishes of a small size are therefore at considerable risk of major uncertainties. When analysing grid squares with extreme emission values, the square containing the maximum livestock emission belonged to a 3 km^2 parish, and the square containing the maximum fertilizer emission belonged to a 2 km^2 parish. It may be concluded that small parishes are at larger risk of over- and underestimation of emissions. It is recommended that the smallest parishes should be aggregated with neighbouring parishes, in order to minimize the risk of these types of errors. It is however difficult to identify a threshold where parishes should be aggregated, as this depends on the area per holding, which in turn varies regionally (see Table 2). A comparison with agricultural land within each parish indicates that 8 km^2, may be a suitable threshold for minimum usable parish size (see Section 3.4 and Figure 5).

An alternative approach to minimize the error is to aggregate the final emission map to a coarser resolution. This is the method typically applied to UK ammonia emission maps, where the final map is normally aggregated to a 5-km grid resolution, both to reduce some of the uncertainties due to the MAUP, and to preserve confidentiality of individual holdings (Dragosits *et al.*, 1998, Hellsten *et al.*, 2008).

Table 2

3.3 Frequency distributions

Figure 4 shows the frequency distribution for each aggregation level, i.e. number of grid cells (1 x 1 km) belonging to each emission class (1 kg NH_3-N ha^{-1} yr^{-1}). The frequency distributions are quite similar for the 5-km, 10-km and parish aggregation. The emission distribution based on the 1-km aggregation level is associated with a higher number of grid cells with low emission values (< 4 kg NH_3-N ha^{-1} yr^{-1}) compared with the other aggregation levels. Also the number of grid cells with high emission values are more frequent in the 1 km level aggregation. 37 grid cells have an emission value which is higher than 1000 kg NH_3-N ha^{-1} yr^{-1} compared with 4-8 grid cells for the other distributions. This clearly shows that the magnitude and number of extreme values are reduced with larger aggregation zones (the smoothing effect). The variation between the frequency distributions is a result of both the scaling and the zonation effect. The scaling effect mainly impacts the magnitude of extreme values, while variations in the spatial distribution of extreme values are explained by the zonation effect.

Figure 4

3.4 Overestimated grid cells regarding agricultural land

Discrepancies between the survey data and the land cover map were also assessed. Table 3 shows the percentage of grid cells (1 km x 1 km) in England where the agricultural area exceeds 100 % and 110 % of the area of the aggregation zone at each aggregation level, i.e. the agricultural land area (crops and grass) of all agricultural holdings in a zone exceed the extent of some of the grid cells (1 x 1 km), after the redistribution within the aggregation zone (based on landuse data). Only 2.5 % of the grid squares exceed the areal extent of the aggregation zone at the 10-km level, with 4.1 % at the 5-km level and as many as 34.5 % at the 1-km level. In the parish distribution, 14.6 % of the squares were overestimated, and consequently others (not identifiable, but presumably nearby), were underestimated. From these values it is clear that the smoothing effect of larger aggregation zones reduces the number of overestimated squares.

Table 3

Figure 5a shows a graph of the ratio of agricultural area within each parish plotted against the size of the parish. The plot shows huge over estimations of agricultural land within the smaller parishes. The graph could be used to select a threshold for minimum usable parish size. Figure 5b indicates that parishes of 7 km^2 or smaller may be associated with large uncertainties due to the overestimation of agricultural land (300 % or more). Hence, 8 km^2 parishes may be used as a threshold minimum ideal parish size for ammonia emission modelling.

Figure 5

3.6 Comparison with measurements

The modelled annual average concentration field for ammonia (1 x 1 km) generated with the FRAME model is shown in Figure 6. Ammonia emissions are based on the 1-km aggregation level for England, and the parish aggregation level for the rest of the UK. Correlation plots comparing annually averaged ammonia concentration measurements from the NAMN for 2002 with modelled ammonia concentrations are shown in Figure 7 (all 91 sites) and Figure 8 (only nature reserve and semi-natural grassland sites, 43 sites). There is more scatter using all sites due to the high local scale variability of NH_3 concentrations in a mixed agricultural landscape. For the nature reserve and semi-natural grassland sites (Figure 8) there is better correlation (higher R^2) but a higher NMB (Normalised Mean Bias) due to small nature reserves being located in a model grid square with significant NH_3 emissions. There does not appear to be much difference in the R^2 between simulations 1-4 for the nature reserve and semi-natural grassland sites, but there is a lower NMB for the simulation plot with 1 km resolution emissions due to less spatial overlap from surrounding agricultural emissions.

Figure 6

Figure 7

Figure 8

3.7 Evaluating the results

Many of the methods applied in this study to evaluate the effect of the Modifiable Areal Unit Problem have clearly shown the smoothing effect that occurs with increased sizes of aggregation zones. This smoothing effect has the advantage of making the result more statistically stable, but a key disadvantage is that spatial details are lost.

Cockings and Martin (2005) suggest that purpose specific automatically designed aggregations may be more appropriate that pre-existing zoning systems. It is not straightforward to conclude what zone size is preferable for the data in this study, if indeed there is a single most suitable value, or whether this may vary regionally. There are probably different ideal zone sizes/shapes vs. farm size ratios across the country, depending on the nature of farming activities and the landscape/geography. The answer is in the relative size of the parishes and the farms across the landscape, the range of farm sizes in a local area, as well as how compact the area of each farm is i.e., whether all the fields are in a neat cluster, or to what extent they are interspersed with neighbouring farmers' fields.

The parish data contain a large range of zonal sizes, with small parishes containing (sometimes) too much spatial detail, while larger parishes are more statistically stable. Some of the extreme and unrealistic values resulting for many very small parishes can be remedied by aggregating the smallest parishes (e.g. 7 km^2 or smaller) with a neighbouring parish. Extreme values as a result of a limited amount of land cover types suitable for specific emission sources within a parish are more difficult to pinpoint and remedy.

The choice of zonal system is limited by data availability, and for some parts of the UK, parish-based agricultural statistics are the only option available. Furthermore, there is a danger in using a zonal system that is different from the parish zones, as some of the holdings are placed randomly within the associated parish by the government statisticians, in the absence of detailed geo-referenced location data. In the parish zone approach, these farms are at least allocated to the correct parish.

3.8 Effects of the MAUP on spatial ammonia emission inventories

This study has shown that the MAUP can have a significant effect on the location of emissions in spatial inventories where aggregated, zonal data are used. If the aggregation zones are small, there is a high risk of the MAUP artificially intensifying emissions from area sources such as agriculture, resulting in unrealistic emission peaks or "hot spots". These peaks are inevitably propagated when the spatial inventories are used as input to atmospheric dispersion and deposition models, thereby overestimating the concentrations or deposition of ammonia in some areas, and underestimating in other areas. These errors may consequently highlight areas that exceed the critical level for ammonia or critical load for nitrogen or acidity, or conversely underestimate effects on sensitive vegetation elsewhere. In principle corrections for this error could be made, but would require assumptions as to the necessary land requirements for different activities over which the particular farms are really using (Hellsten, 2006).

If the aggregation zones on the other hand are too large, true emission peaks are likely to be smoothed out, hence underestimating areas exceeding the critical levels or loads. Spatial accuracy and reducing spatial uncertainties are therefore important tasks to be included in emission inventory compilation. It is recognised that it is impossible to eliminate totally the effects of the MAUP, but raising the awareness of these types of issues, especially among people working with zonal data as inputs to spatial emission inventories, may improve understanding and interpretation of the results. Flowerdew (2011) concluded that the problem of the MAUP is not very serious most of the time, but for certain occasions it can make a great difference. The difficulty lies is predicting these occasions.

When dealing with aggregated point data, the Modifiable Areal Unit Problem will always be present, unless the spatial distribution is totally homogenous, which is unlikely for most distributions. It is certainly not the case when dealing with parish-aggregated agricultural survey data, as these are based on historic and often artificial administrative borders.

We suggest the following methods to mitigate the effects of the Modifiable Areal Unit Problem in the context of spatial ammonia emission inventories:

- Aggregating small parishes (≤ 7 km^2) with neighbouring parishes.
- Incorporating additional data (e.g., land cover data) to distribute the survey items as emission sources within each zone.
- Dealing with large point sources such as pig and poultry production facilities in a particular manner, e.g.
    - By using additional data (remote sensing data or data from the IPPC-database) to identify these facilities.
    - By applying assumptions as to the necessary land requirements for pig and poultry manure. For instance, an iterative process could be applied, where pig and poultry manure in aggregation zones with insufficient land suitable for manure spreading is transported further away from the farm, as suggested by Hellsten (2006).

No matter which approach that is used for the redistribution of agricultural statistics within the landscape, we are convinced that having agricultural statistics made available at a fine resolution would improve both the possibility to analyse the data, and to better understand the problems and effects of the Modifiable Areal Unit Problem on spatial ammonia emission inventories.

4. Acknowledgements

This work was funded by Defra, through the NARSES project (Project No AM0101), the NERC ASSIST project and through the National Atmospheric Emission Inventory (NAEI) and with underpinning science funding from the NERC Centre for Ecology & Hydrology (CEH). The authors are grateful to Defra for providing data for this study.

5. References

Asman W.A.H., Sutton M.A. and Schjorring J.K. 1998. Ammonia: emission, atmospheric transport and deposition. New Phytologist 139, 27-48.

Beusen A.H.W., Bouwman A.F., Heuberger P.S.C., Van Drecht G. and Van Der Hoek K.W. 2008. Bottom-up uncertainty estimates of global ammonia emissions from global agricultural production systems. Atmospheric Environment, V 42, I 24, 6067-6077.

Cellier P., Durand P., Hutchings N., Dragosits U., Theobald M.R., Drouet J.L., Oenema O., Bleeker A., Breuer L., Dalgaard T., Duretz S., Kros J., Loubet B., Olesen J.E., Mérot P., Viaud V., de Vries W. and Sutton M.A. 2011. Dispersion and fate of nitrogen in rural landscapes. Chapter 11. In: Sutton M.A., Howard C.M., Erisman J.W., Billen G., Bleeker A., Grennfelt P., van Grinsven H. and Grizzetti B. (eds) The European Nitrogen Assessment. Cambridge University Press. p.229-248.

Cockings S. and Martin D. 2005. Zone design for environment and health studies using pre-aggregated data, Social Science & Medicine, Volume 60, Issue 12, June 2005, 2729-2742.

Dai J. and Rocke D. 2000. Modeling spatial variation in area source emissions, Journal of Agricultural, Biological, and Environmental Statistics, Vol. 5, No. 1, 7-21.

Dark S.J. and Bram D. 2007. The modifiable areal unit problem (MAUP) in physical geography, Progress in Physical Geography 31 (5), 471-479.

Dore A. J., Vieno M., Tang Y. S., Dragosits U., Dosio A., Weston K. J. and Sutton M. A. 2007. Modelling the atmospheric transport and deposition of sulphur and nitrogen over the United Kingdom and assessment of the influence of SO_2 emissions from international shipping. Atmospheric Environment, 41 (11). 2355-2367.

Dore A.J., Kryza M., Hall J. Hallsworth S., Keller V., Vieno M. & Sutton M.A. (2012) The Influence of Model Grid Resolution on Estimation of National Scale Nitrogen Deposition and Exceedance of Critical Loads Biogeosciences 9, 1597-1609

Dore A.J., D. Carslaw, C. Braban, M.Cain, C. Chemel, C. Conolly, R.G. Derwent, S.J. Griffiths, J. Hall, G. Hayman, S. Lawrence, S.E. Metcalfe, A. Redington, D. Simpson, M.A. Sutton, P. Sutton, Y.S. Tang, M. Vieno M. Werner J.D. Whyatt (2015) Evaluation of the performance of different atmospheric chemical transport models and inter-comparison of nitrogen and sulphur deposition estimates for the UK. Atmospheric Environment 119, 131-143

Dragosits U., Sutton M.A., Place C.J. and Bayley A.A. 1998. Modelling the spatial distribution of agricultural ammonia emissions in the UK. Environmental Pullution 102, 195-203.

Dragosits U., Theobald M.R., Place C.J., ApSimon H.M. and Sutton M.A. 2006. Spatial planning at the landscape level to mitigate the effects of atmospheric ammonia deposition. Journal of Environmental Science and Policy 9, 626-638.

Dragosits U., Theobald M.R., Place C.J., Lord E., Webb J., Hill J., ApSimon H.M. and Sutton M.A. 2002. Ammonia emission, deposition and impact assessment at the field scale: a case study of sub-grid spatial variability. Environmental Pullution 117, 147-158.

EU 2003. Regulation (EC) No 1059/2003 of the European Parliament and of the Council of 26 May 2003 on the establishment of a common classification of territorial units for statistics (NUTS)

Flowerdew R. 2011. How serious is the Modifiable Areal Unit Problem for analysis of English census data? Population Trends 145, Autumn 2011, Office for National Statistics, 106-118.

Geddes A., Gimona A. and Elston D.A. 2003. Estimating local variations in land use statistics. International Journal of Geographical Information Science 17, 299-319.

Geels C., Brandt J., Christensen J.H., Frohn L.M., and Hansen K.M. 2005. Long-Term Calculations with a Comprehensive Nested Hemispheric Air Pollution Transport Model. In: Faragó, I., Georgiev, K. and Havasi, Á. (eds) Advances in Air Pollution Modeling for Environmental Security, Springer Netherlands, 185-196.

Hallsworth S., Dore A.J., Bealey W.J., Dragosits U., Vieno M., Hellsten S., Tang Y.S. and Sutton M.A. 2010. The role of indicator choice in quantifying the threat of atmospheric ammonia to the 'Natura 2000' network. Environmental Science & Policy 13, 671-687.

Hellsten S. 2006. Improving the spatial distribution of ammonia emissions from poultry farming in the UK, Chapter 7 in A spatio- temporal ammonia emissions inventory for the UK: Thesis (Ph.D.) University of Edinburgh, Scotland, UK, 2006. https://www.era.lib.ed.ac.uk/handle/1842/24693.

Hellsten S., Dragosits U., Place C.J., Misselbrook T.H., Tang Y.S. and Sutton M.A. 2007. Modelling seasonal dynamics from temporal variation in agricultural practices in the UK ammonia emission inventory. Water Air and Soil Pollution, Focus, Volume 7, Issues 1-3, 3-13.

Hellsten S., Dragosits U., Place C.J., Vieno M., Dore A.J., Misselbrook T.H., Tang Y.S. and Sutton M.A. 2008. Modelling the spatial distribution of ammonia emissions in the UK, Environmental Pollution 154, 370-379.

Jelinski D.E. and Wu J.G. 1996. The modifiable areal unit problem and implications for landscape ecology. Landscape Ecology 11, 129-140.

Kühlwein J. and Friedrich R. 2000. Uncertainties of modelling emissions from road transport. Atmospheric Environment 34, 4603-4610.

Leopold U., Heuvelink G.B.M, Drouet L. and Zachary D.S. 2012. Modelling Spatial Uncertainties associated with Emission Disaggregation in an integrated Energy Air Quality Assessment Model, International Environmental Modelling and Software Society (iEMSs) 2012 International Congress on Environmental Modelling and Software Managing Resources of a Limited Planet, Sixth Biennial Meeting, Leipzig, Germany R. Seppelt, A.A. Voinov, S. Lange, D. Bankamp (Eds.) http://www.iemss.org/society/index.php/iemss-2012-proceedings

Lindley S.J., Conlan D.E., Raper D.W. and Watson A.F.R. 2000. Uncertainties in the compilation of spatially resolved emission inventories - evidence from a comparative study. Atmospheric Environment 34, 375-388.

Longley P. and Batty M. 1996. Analysis, modelling forecasting, and GIS technology, in: Longley P, Batty M. (Eds.), Spatial Analysis: modelling in a GIS environment. Geoinformation International. Cambridge, pp 1-15.

Loubet B., Asman W.A.H., Theobald M.R., Hertel O., Tang Y.S., Robin P., Hassouna M., Dämmgen U., Genermont S., Cellier P. and Sutton M.A. 2009. Ammonia Deposition near Hot Spots: Processes, Models and Monitoring Methods. In Sutton, M.A., Reis, S. and Baker, S.M.H. (eds) Atmospheric Ammonia. Detecting Emissions Changes and Environmental Impacts, Springer, 205-267.

Maes J., Vliegen J., Van de Vel K., Janssen S., Deutsch F., De Ridder K., Mensink C. 2009. Spatial surrogates for the disaggregation of CORINAIR emission inventories, Atmospheric Environment, Volume 43, Issue 6, February 2009, 1246-1254.

Matejko M., Dore A.J., Hall J., Dore C.J., Błaś M., Kryza M., Smith R. and Fowler D. 2009. The influence of long term trends in pollutant emissions on deposition of sulphur and nitrogen and exceedance of critical loads in the United Kingdom, Environmental Science & Policy 12, 882 – 896.

Misselbrook T.H., Van der Weerden T.J., Pain B.F., Jarvis S.C., Chambers B.J., Smith K.A., Phillips V.R. and Demmers T.G.M. 2000. Ammonia emission factors for UK agriculture. Atmospheric Environment 34, 871-880.

Montello D.R. 2001. Scale in Geography, in: N.J. Smelser, P.B. Baltes (Eds.), International Encyclopedia of the Social & Behavioral Sciences, Pergamon Press, Oxford, pp. 13501-13504.

Openshaw S. 1977. A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling. Transactions of the institute of British Geographers, New series 2, 459-472.

Openshaw S. 1984. The Modifiable Areal Unit Problem: Concepts and techniques in Modern Geography 38, Geobooks, Norwich.

Openshaw S. and Alvanides S. 1999. Designing Zoning Systems for representation of socio-economic data. in I. Frank *et al.* (Eds.), Time and Motion of Socio-Economic Units, GISDATA Series, Taylor and Francis, London.

Openshaw S. and Rao L. 1995. Re-engineering 1991 census geography: serial and parallel algorithms for unconstrained zone design, Research Paper 95-3, School of Geography, Leeds University, Leeds.

Openshaw S. and Taylor P. 1979. A million or so correlation coefficients: Three experiments on the modifiable area unit problem, in: N. Wrigley (Eds.), Statistical Applications in the Spatial Sciences, Pion, London, pp. 127-144.

Sadahiro Y. 2000. Accuracy of count data transferred through the areal weighting interpolation method. International Journal of Geographical Information Science 14, 25-50.

Sutton M.A., Milford C., Dragosits U., Place C.J., Singles R.J., Smith R.I., Pitcairn C.E.R., Fowler D., Hill J., ApSimon H.M., Ross C., Hill R., Jarvis S.C., Pain B.F., Phillips V.C., Harrison R., Moss D., Webb J., Espenhahn S.E., Lee D.S., Hornung M., Ullyett J., Bull K.R., Emmett B.A., Lowe J. and Wyers G.P., 1998. Dispersion, deposition and impacts of atmospheric ammonia: quantifying local budgets and spatial variability. Environmental Pollution 102, 349-361.

Sutton M.A., Place C.J., Eager M., Fowler D. and Smith R.I. 1995. Assessment of the magnitude of ammonia emissions in the United-Kingdom. Atmospheric Environment 29, 1393-1411.

Sutton M.A., Miners B., Tang Y.S., Milford C., Wyers G.P., Duyzer J.H. and Fowler D. 2001. Comparison of low costs measurements techniques for long-term monitoring of atmospheric ammonia. Journal of Environmental Monitoring 3, 446-453.

Sutton M.A., Reis R., Riddick S.N., Dragosits U., Nemitz E., Theobald M.R., Tang Y.S., Braban C.F., Vieno M., Dore A.J., Mitchell R.F., Wanless S., Daunt F., Fowler D., Blackall T.D., Milford C., Flechard C.F., Loubet B., Massad R., Cellier P., Coheur P.F., Clarisse L., van Damme M., Ngadi Y., Clerbaux C., Skjøth C.A., Geels C., Hertel O., Wichink Kruit R.J., Pinder R.W., Bash J.O., Walker J.D., Simpson D., Horvath L., Misselbrook T.H., Bleeker A., Dentener F. & Wim de Vries V. 2013. Toward a climate-dependent paradigm of ammonia emission & deposition. *Phil. Trans. Roy. Soc. (Ser. B).* 368: 20130166. http://dx.doi.org/10.1098/rstb.2013.0166

Tang Y.S., Cape J.N. and Sutton, M.A. 2001. Development and types of passive samplers for monitoring atmospheric NO_2 and NH_3 concentrations. In Proceedings of the International Symposium on Passive Sampling of Gaseous Air Pollutants in Ecological Effects Research. The Scientific World, 1, 513-529.

Tang Y.S. and Sutton M.A. 2004. Quality Management in the UK National Ammonia Monitoring Network: Proceedings of the International Conference "QA/QC in the field of emission and Air Quality Measurements" 21-23 May 2003, Prague (CZ), EUR 20973 (ISBN 92-894-6523-9).

Tang Y.S. Braban C.F., Dragosits U., Dore A.J., Simmons I., van Dijk N., Poskitt J., Pereira G.D.S., Keenan P.O., Conolly C., Vincent K., Smith R.I., Heal M.R. & Sutton M.A. 2018. Drivers for spatial, temporal and long-term trends in atmospheric ammonia and ammonium in the UK. Atmospheric Chemistry and Physics, 18, 705-733. https://doi.org/10.5194/acp-18-705-2018.

Vieno M. 2006. The use of an Atmospheric Chemistry-Transport Model (FRAME) over the UK and the development of its numerical and physical schemes, Thesis (Ph.D.), University of Edinburgh, Scotland.

Winiwarter W., Dore C., Hayman G., Vlachogiannis D., Gounaris N., Bartzis J., Ekstrand S., Tamponi M. and Maffeis G. 2003. Methods for comparing gridded inventories of atmospheric emissions-application for Milan province, Italy and the Greater Athens Area, Greece. Science of the Total Environment 303, 231-243.

Winiwarter W. and Rypdal K. 2001. Assessing the uncertainty associated with national greenhouse gas emission inventories: a case study for Austria. Atmospheric Environment 35, 5425-5440.

Xu P., Liao Y.J., Lin Y.H., Zhao C.X., Yan C.H. Cao M.N., Wang G.S. and Luan S.J. 2016. High-resolution inventory of ammonia emissions from agricultural fertilizer in China from 1978 to 2008, Atmos. Chem. Phys., 16, 1207-1218, doi:10.5194/acp-16-1207-2016.

Zheng J.Y., Yin S.S., Kang D.W., Che W.W. and Zhong L.J. 2012. Development and uncertainty analysis of a high-resolution NH_3 emissions inventory and its implications with precipitation over the Pearl River Delta region, China, Atmospheric Chemistry and Physics, 12, 7041-7058.

Figures & Figure Captions

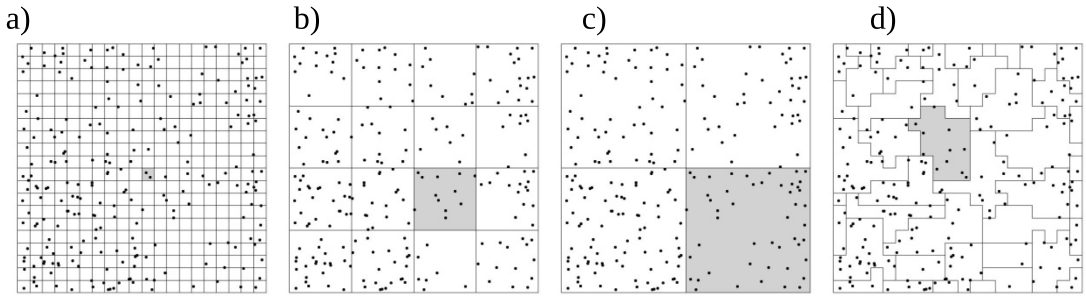


Figure 1. Holding data were aggregated using four different zonal systems: a) 1-km grid, b) 5-km grid, c) 10-km grid, and d) parish zones. The highlighted zones in each figure show the potential area to which emission may be assigned at each zonation level for a farm located in the shaded 1-km grid cell of figure a (also see Figure 2). Area: 400 km^2.

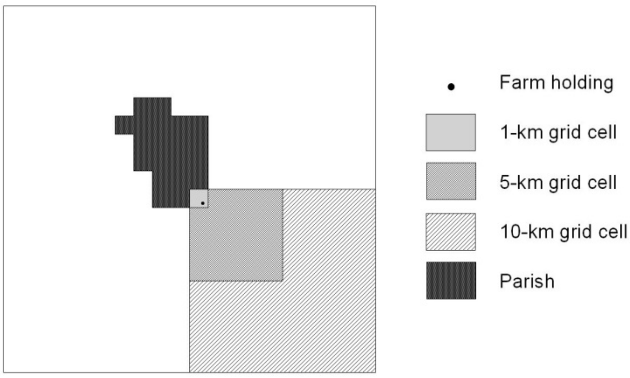


Figure 2. Potential area for re-distribution of emissions from a single farm at different aggregation zones *(*10-km grid, 5-km grid, 1-km grid and parish zones). Both the scale and the zonation effect are demonstrated. Area: 400 km^2.

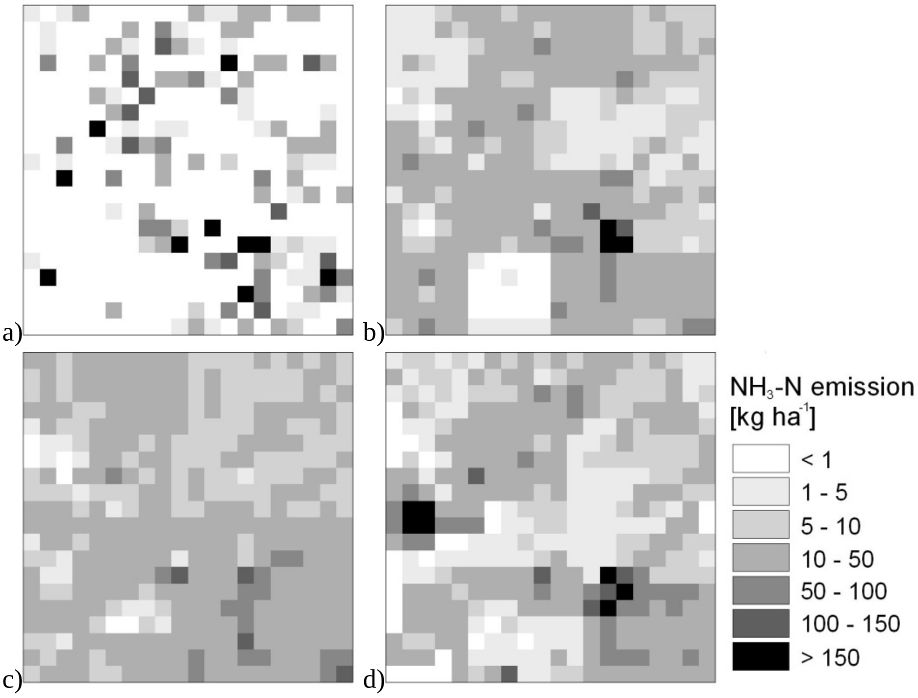


Figure 3. Examples of emission results, based on real data (livestock emissions) at a local scale (area 400 km^2) from a part of England using different aggregation zones: a) 1-km zone, b) 5-km zone, c) 10-km zone and d) parish zone.

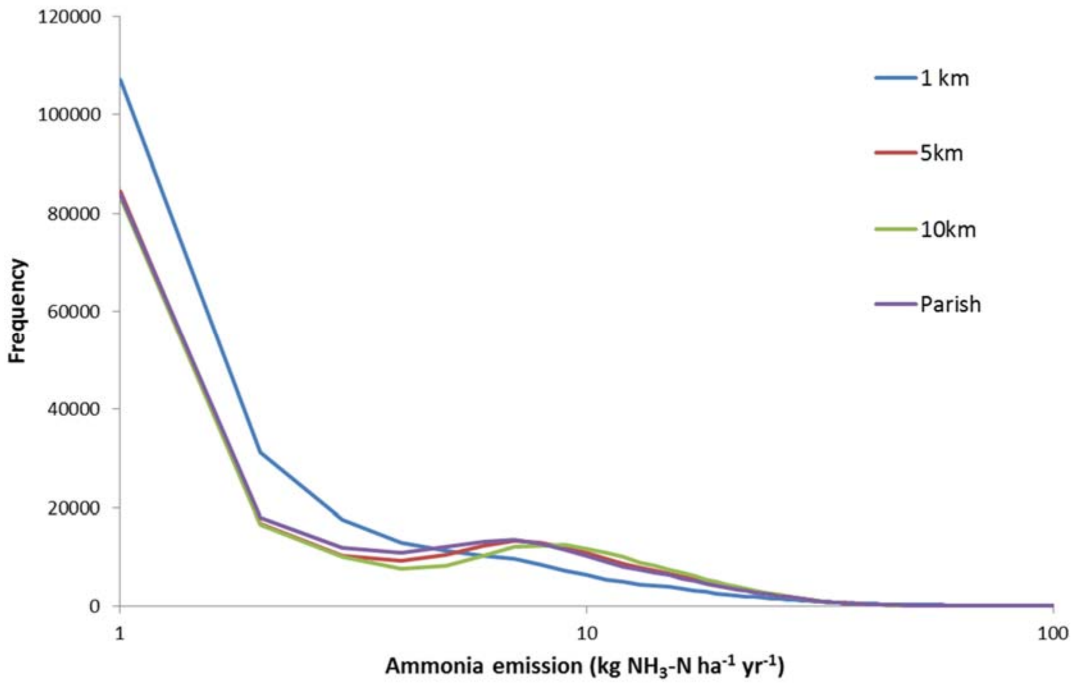


Figure 4. Frequency distribution of the ammonia emissions for the different aggregation levels. The frequency represents number of grid cells (1 x 1 km) within each emission class (1 kg NH_3-N ha^{-1} yr^{-1}).

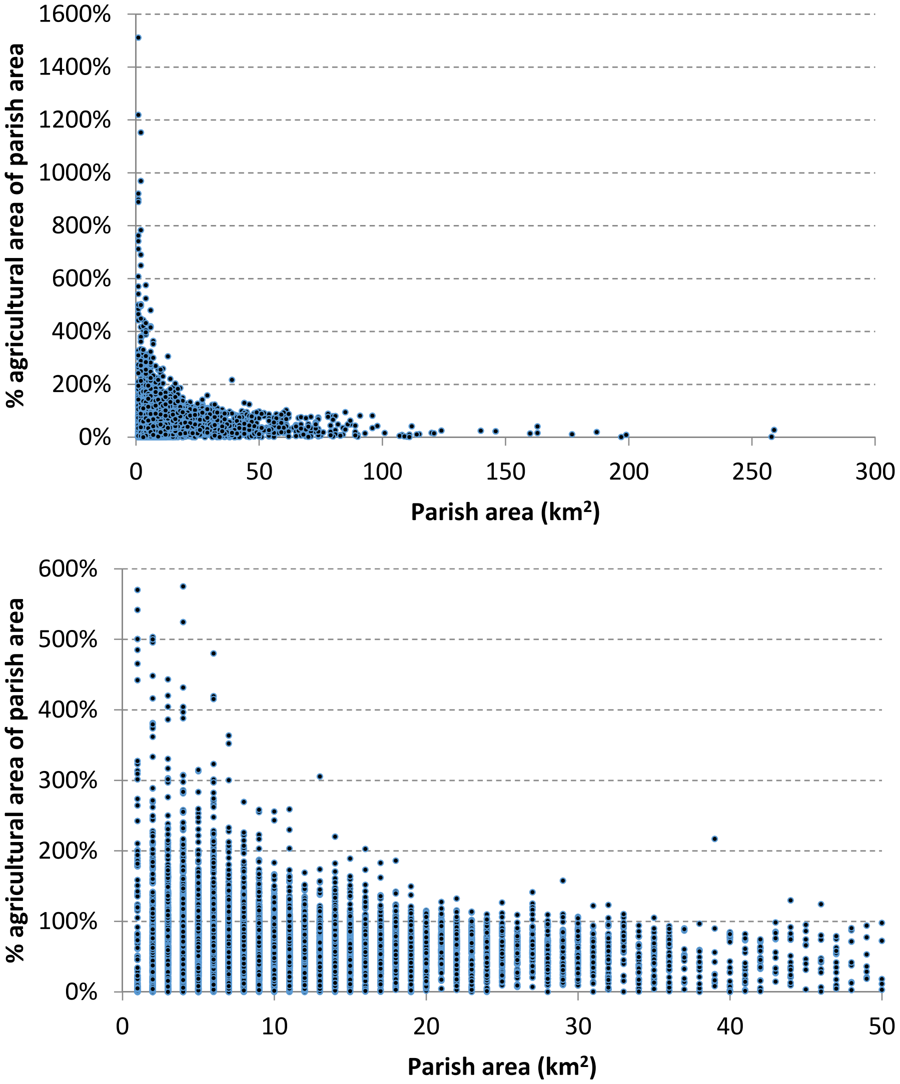


Figure 5. The ratio of agricultural area within each parish plotted against the size of the parish.

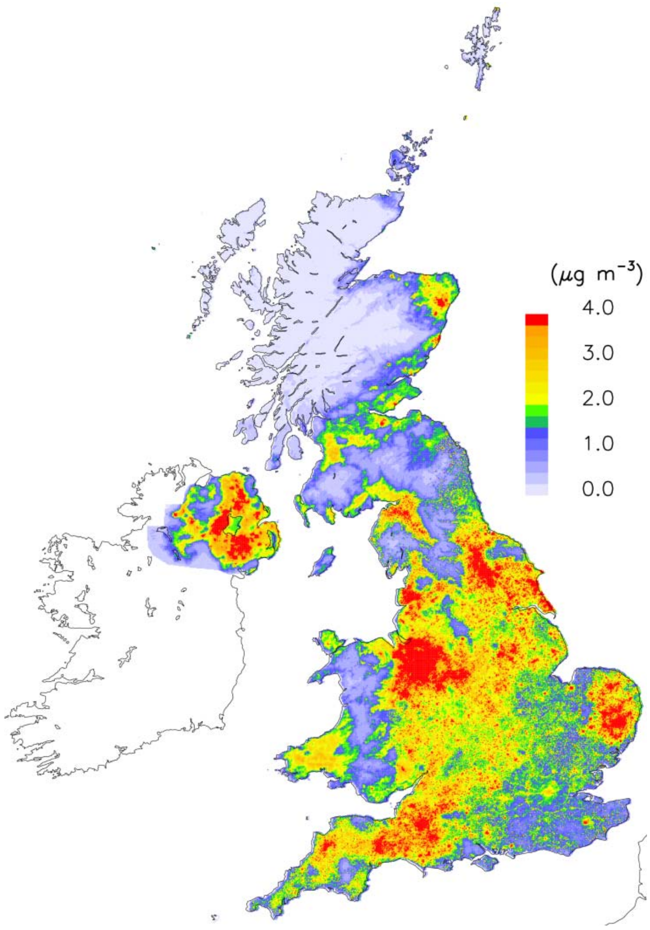
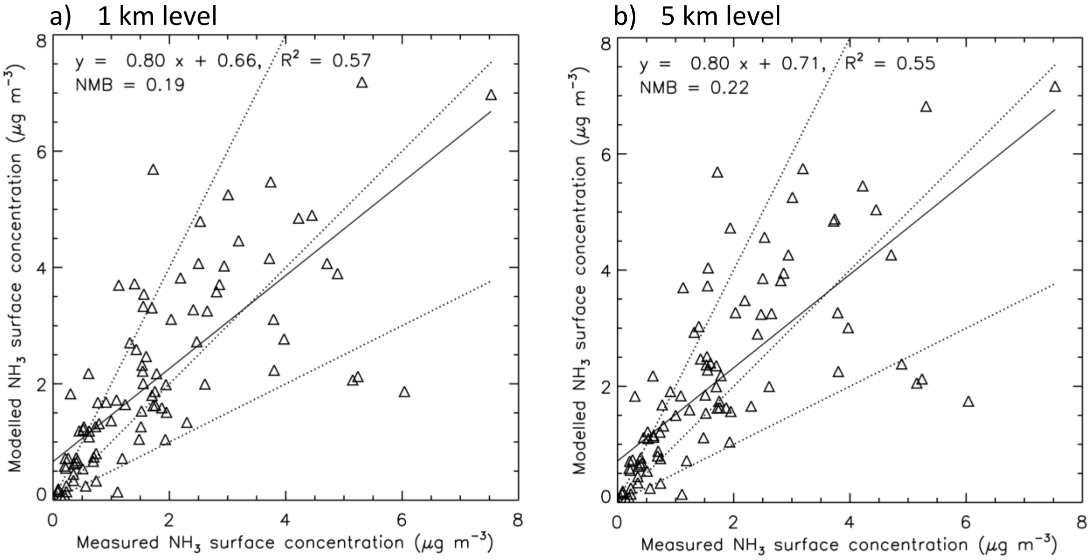


Figure 6. Modelled concentration field for ammonia (1 km x 1 km) with the FRAME model, based on ammonia emission estimates from the 1-km aggregation level in England. The parish distribution was used for the rest of the UK.

a) 1 km level

b) 5 km level



c) 10 km level

d) parish level

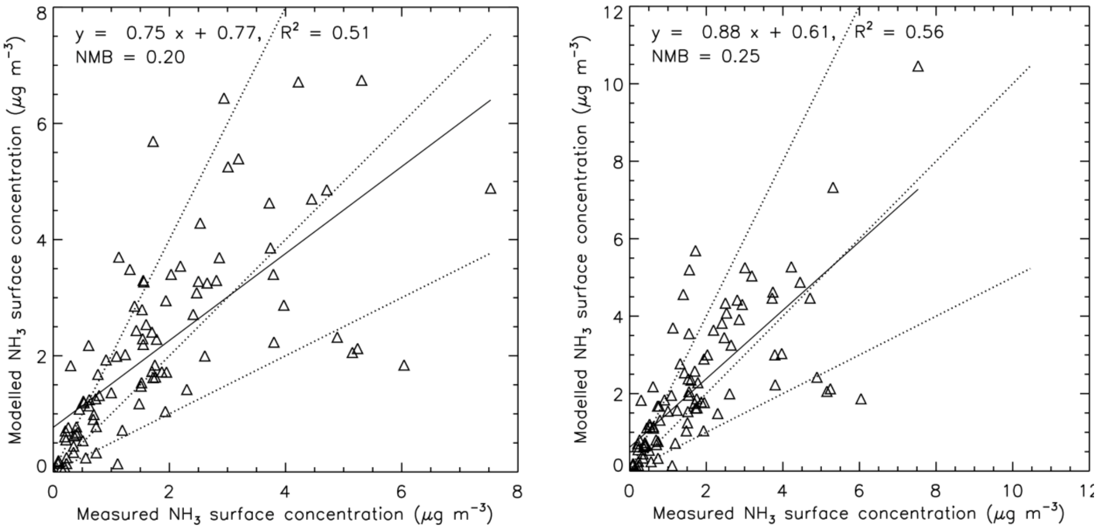
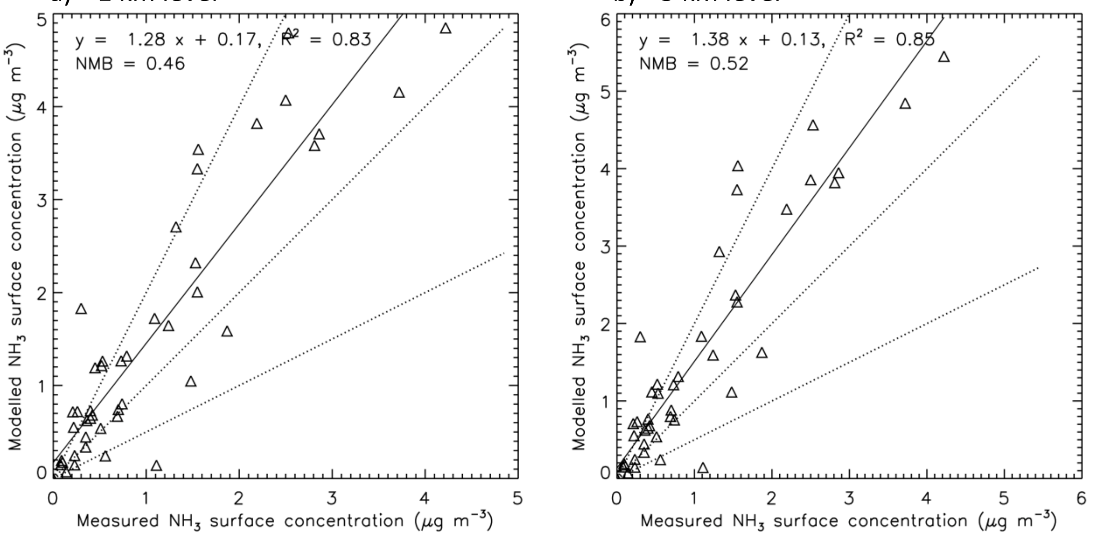


Figure 7. Correlation plots, comparing all annually averaged ammonia measurement results for 2002 (91 sites) with modelled ammonia concentrations based on zone level a) 1 km, b) 5 km, c) 10 km and d) parish level.

a) 1 km level b) 5 km level



c) 10 km level d) parish level

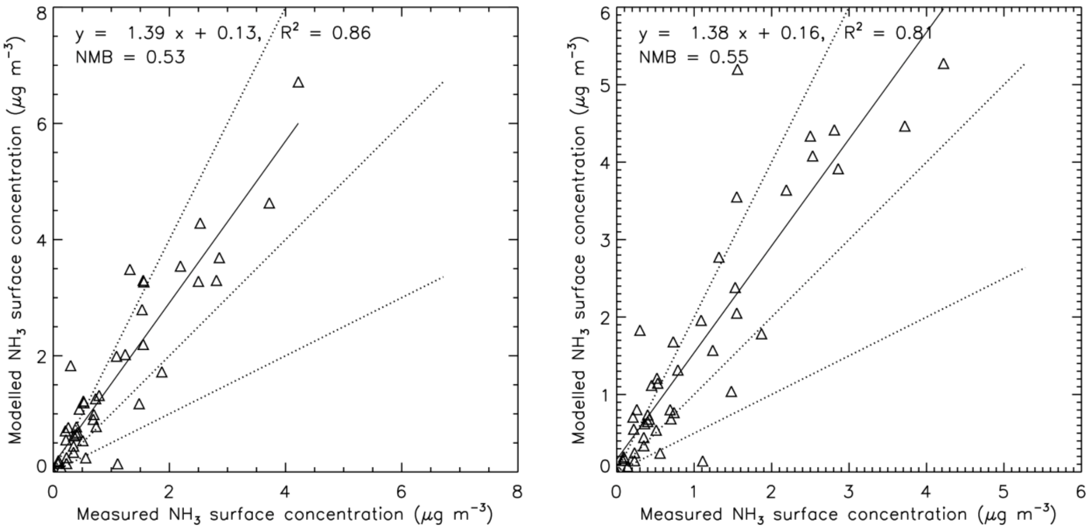


Figure 8. Correlation plots, comparing only nature reserve and semi-natural grassland sites (43 sites, annually averaged ammonia measurement results for 2002) with modelled ammonia concentrations based on zone level a) 1 km, b) 5 km, c) 10 km and d) parish level.

Tables

Table 1. Maximum NH_3-N emission (tonnes/year) per 1-km grid square in the UK-wide emission maps (See extended Figure 5) derived using different zone systems.

Emission (t NH_3 $year^{-1}$)	Aggr.zone: 1-km grid	Aggr.zone: 5-km grid	Aggr.zone: 10-km grid	Aggr.zone: Parish
Livestock	917.4	69.4	34.8	552.5
Fertilizers	14.3	1.2	0.68	5.8

Table 2. Statistics of the English parish dataset, showing various parish size groups, and number of parishes belonging to each size group, for all English parishes (average size 11.8 km^2) on the left, and details for the smallest size group on the right .

Parish size (km^2)	Number of parishes
1-5	2, 940
6-25	7, 338
26-100	818
> 100	25
Total	11,121

Parish size	No of parishes
1 km^2	269
2 km^2	349
3 km^2	555
4 km^2	846
5 km^2	921

Table 3. Percentage of grid squares where the area of agricultural land is overestimated by aggregating holdings to zone systems, i.e., grid cells where the sum of agricultural area from all holdings allocated to the zone is larger than the area of the grid cell itself (crops & grass > 100 and 110 ha).

Zone level	> 100 %	> 110 %
1-km grid	34.5 %	31.3 %
5-km grid	4.1 %	2.3 %
10-km grid	2.5 %	1.0 %
Parish	14.6 %	9.7 %