Environmental Pollution 240 (2018) 412–421

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

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

A R T I C L E I N F O

Article history:
Received 11 September 2017
Received in revised form
27 April 2018
Accepted 27 April 2018

Keywords:
Modifiable areal unit problem
Zone design
Aggregated data
Ammonia
Agricultural survey data
Spatial inventories

A B S T R A C T

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.

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).

* Corresponding author. IVL Swedish Environmental Research Institute Ltd, P.O. Box 5302, SE-400 14 Gothenburg, Sweden.
E-mail addresses: sofie.hellsten@ivl.se (S. Hellsten), ud@ceh.ac.uk (U. Dragosits), cjp@staffmail.ed.ac.uk (C.J. Place), todo@ceh.ac.uk (A.J. Dore), yst@ceh.ac.uk (Y.S. Tang), ms@ceh.ac.uk (M.A. Sutton).

https://doi.org/10.1016/j.envpol.2018.04.132
0269-7491/

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, 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. (2005) 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 Alvanides, 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 × 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 (Fig. 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.