An analysis of the likely success of policy actions under uncertainty: recovery from acidification across Great Britain

J. Duncan Whyatt[a], Sarah E. Metcalfe*[,b], Richard G. Derwent[c], Trevor Page[a]

[a]Lancaster Environment Centre, Lancaster University, LA1 4YQ, United Kingdom, d.whyatt@lancaster.ac.uk; t.page@lancaster.ac.uk

[b]School of Geography, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom, sarah.metcalfe@nottingham.ac.uk

[c]rdscientific, Newbury, Berkshire, RG14 6LH, United Kingdom, r.derwent@btopenworld.com

* Corresponding Author: sarah.metcalfe@nottingham.ac.uk; tel.: +44 115 846 7712: fax: +44 115 951 5249

Environmental Science and Policy (in press)

ABSTRACT

In the context of wider debates about the role of uncertainty in environmental science and the development of environmental policy, we use a Generalised Likelihood Uncertainty Estimate (GLUE) approach to address the uncertainty in both acid deposition model predictions and in the sensitivity of the soils to assess the likely success of policy actions to reduce acid deposition damage across Great Britain. A subset of 11, 699 acid deposition model runs that adequately represented observed deposition data were used to provide acid deposition distributions for 2005 and 2020, following a substantial reduction in SO_2 and NO_x emissions. Uncertain critical loads data for soils were then combined with these deposition data to derive estimates of the accumulated exceedance (AE) of critical loads for 2005 and 2020. For the more sensitive soils, the differences in accumulated exceedance between 2005 and 2020 were such that we could be sure that they were significant and a meaningful environmental improvement would result. For the least sensitive soils, critical loads were largely met by 2020, hence uncertainties in the differences in accumulated exceedance were of little policy relevance. Our approach of combining estimates of uncertainty in both a pollution model and an effects model, shows that even taking these combined uncertainties into account, policy-makers can be sure that the substantial planned reduction in acidic emissions will reduce critical loads exceedances. The use of accumulated exceedance as a relative measure of environmental protection provides additional information to policy makers in tackling this ‘wicked problem’.

Keywords: HARM, GLUE, uncertainty, critical loads, soil acidification

1. Introduction

The many types of uncertainty that can affect policy making and how these can be presented to and then handled by policy makers, have become topics of increasing interest. Schneider and Kuntz-Duriseti (2002) considered uncertainty in climate change policy. They suggested that whilst one approach is to reduce (bound) the uncertainty by collecting more data, more understanding and building better models, the other approach is to reduce the effects of (manage) any uncertainty in understanding by taking it into account in policy making. This second approach can be traced back to ideas about ecosystem resilience and recovery after disturbance developed in the 1970s. Refsgaard et al. (2007) in a review of uncertainty in the context of water management, suggested that uncertainty in its widest sense can usefully be regarded as the degree of confidence a decision maker has about possible outcomes and/or the probabilities of these outcomes. Uusitalo et al. (2015) suggested that uncertainty analysis can provide decision makers with a realistic picture of possible outcomes, in a context where results are going to be better or worse, not true or false, i.e. that environmental problems are 'wicked problems'. Whilst some types of uncertainty are unquantifiable, other types can be quantified through approaches such as sensitivity analysis, the use of multiple models and exploring the impact of parameter uncertainty. Here we take a quantitative approach to uncertainty in the context of recovery from the problem of acidification in Great Britain. We quantify and then combine the uncertainties in outputs from one acid deposition model and one measure of ecosystem health to assess whether current emissions reduction policies are likely to deliver ecosystem protection. We believe that this is the first effort to combine the uncertainties in both these elements in a single assessment.

European policymakers have been concerned about the acidification of sensitive soils and terrestrial ecosystems, driven by emissions of acidic species, sulphur dioxide (SO_2) and nitrogen oxides (NO_x) since the 1970s. These concerns have led to concerted policy actions within the United Nations Economic Commission for Europe (UN ECE) and the European Union (EU), designed to reduce emissions and hence, the damaging deposition. The UN ECE agreed the Convention on Long-Range Transboundary Air Pollution (CLRTAP) in 1979 and has since promulgated a series of Protocols to the Convention, initially involving SO_2 and NO_x separately and then combined with ammonia (NH_3) under the Gothenburg Protocol (1999), referred to as the 'Multi-pollutant, Multi-effect Protocol'. A revision of the Gothenburg Protocol was agreed in 2012 (referred to here as RGP, see Amann et al., 2012; Reis et al., 2012). The EU has tackled the need to reduce emissions through a series of directives focussing initially on Large Combustion Plant (1988 and 2001), giving rise to the National Emission Ceilings Directive (NECD). In 2005, the EU put forward its Thematic Strategy on Air Pollution, Clean Air for Europe (CAFÉ) and under this framework is renegotiating the NECD with current commitments extending to 2029, with new commitments after 2030 (for an assessment of the NECD see Hettelingh et al., 2013a). Within these policy contexts, the chosen measure of ecosystem sensitivity was the critical load (CL) (Hettelingh et al., 1995), where the CL is the amount of deposition the chosen receptor can apparently tolerate without damage being likely (Bull, 1992). Where deposition was greater than (exceeded) the CL, damage was assumed to occur. CLs have been developed for a range of receptors (soils, freshwaters and a variety of terrestrial ecosystems) using a number of different methodologies (for the latest UK information see http://www.cldm.ceh.ac.uk/, for details of the most recent changes in methodology across Europe see Slootweg et al. 2015). It has been long recognised that there is variability between representations of CLs and that there are uncertainties in their calculation (see Zak et al., 1997), but CLs remain central to policymaking in this area and are an accepted risk assessment tool (Hettelingh et al., 2013b; Holmberg et al., 2013). The success of any emissions reduction policy is gauged by the resulting reduction in CL exceedance and system recovery (chemical and biological) (Posch et al., 2012), recognising that any system is unlikely to recover to exactly its pre-acidification state (Helliwell et al., 2014).

As it soon became evident that CLs would not be achievable across the whole of Europe in the foreseeable future, the concept of 'gap-closure' was adopted to formulate acid deposition policies (see Amann et al., 2012 and the references therein). Gap closure implies reducing CL exceedance by a given fraction, say 50%, and then using integrated assessment modelling to find an equitable and fair distribution of emission reductions across the European countries to achieve the gap-closure target. Whilst this is a pragmatic approach, the approach cannot use meeting CLs as its optimisation target (and hence cannot guarantee complete ecosystem protection) and so a new index of environmental protection has been defined in terms of reducing 'accumulated exceedance' (AE) which captures both the magnitude and areal extent of exceedance. This index requires the combination of both CL and acid deposition data, both of which are uncertain.

The historical reductions in emissions across the EU-28 countries (by 87% for SO_2, 54% for NO_x and 27% for NH_3 since 1990) (European Environment Agency (EEA), 2015) and measured decreases in deposition, have been reflected by measurable recovery in pH and acid neutralising capacity in many surface waters (Battarbee et al., 2014; Kernan et al., 2010) and reductions in CL exceedance (De Wit et al., 2015; RoTAP, 2012). Forward projections of current emission reduction commitments and the agreement of any additional reductions, however, depend on the application of atmospheric transport and deposition models, whose outputs can then be compared with CLs to assess the likely resulting environmental improvement (gains). Acid deposition models are uncertain because the parameterisations on which they are based and the input parameters that are fed into them, both contain simplifications and assumptions. CL are also uncertain, as described above. It is important, therefore, that policymakers have confidence in the outcomes of this modelling procedure (deposition and CL exceedance) given all the uncertainties inherent in both the atmospheric transport and CL models and can be assured that the higher costs of additional future emission reductions (assuming that the cheaper options have already been adopted) will actually increase protection of sensitive ecosystems and that recovery from acidification will continue. Two questions therefore arise: 1) can we can really be sure that the emissions reductions proposed to reduce AE will produce discernible environmental improvement or will they be lost in uncertainty? and 2) does the change of approach from an absolute target (CL exceeded or not) to a relative one (based on accumulated exceedance), change our perception of environmental improvement? Here we address both these questions. The concerns around the implications of scientific and model uncertainty for policy making that we address here in relation to acidification are relevant across a range of environmental issues.

We address our two questions about the impact of scientific uncertainty on achieving environmental protection, by exploring the impact of uncertainties in one atmospheric transport and deposition model, the Hull Acid Rain Model (HARM, Metcalfe et al., 2005) and one representation of CL (for soils), based on the Skokloster classification, by comparing estimates of accumulated exceedance of CL in 2005 and 2020 and assessing the likelihood of environmental protection across Great Britain (GB). This builds on an initial assessment of the impacts of uncertainty in HARM on CL exceedance across Wales reported by Heywood et al. (2006a). We provide a brief description of HARM and set out our approach to representing uncertainty in HARM and the CL for soils data set. We describe how we have combined estimates of deposition and sensitivity to acidification (CLs) to yield estimates of accumulated exceedance (AE) and how we have assessed the significance of the modelled changes. Our method is illustrated with reference to one 10 km x 10 km grid square in the Peak District in northern England, before going on to present and discuss the results for the whole of GB and consider the wider implications of this more rigorous approach for policy making.

2. Methodology

2.1 HARM and the GLUE framework

HARM is a receptor-orientated Lagrangian statistical model which is driven by emissions of SO_2, NO_x and NH_3 across the UK and the wider European area. Over a number of years, the model has been used to help in the formulation of acidification control policies in the UK. It provides estimates of wet and dry sulphur and nitrogen (both oxidised and reduced) depositions at 10 km x 10 km spatial resolution across the UK. Further details of the model are given elsewhere (Dore et al., 2015;

Metcalfe et al., 2005; Whyatt et al., 2007). Here, HARM has been run using 2005 emissions estimates for SO_2, NO_x and NH_3 sources within the UK and the rest of Europe. An illustrative, gap closure type, scenario was then applied to simulate a possible 2020 emission situation involving a 35% reduction in SO_2 emissions and a 33% reduction in NO_x emissions (no reduction was applied to NH_3 emissions). This 2020 scenario was developed before the RGP was agreed, but is broadly consistent with the UK's current Gothenburg commitments (DEFRA, 2015). Our SO_2 emissions lie within the likely ranges for 2020, but our NO_x emissions are a little high. It is also proposed that UK NH_3 emissions will decline by 2020, by around 12% from the figure used here. Because our results are likely to be influenced by the absolute magnitude of the deposition reduction as well as the spatial distribution of any reduction, our illustrative or hypothetical reduction should be within the bounds of current projections.

Policymakers require that any model used for environmental policy formulation should reproduce real world behaviour adequately. In the present context, this means that an acid deposition model should reproduce the observed acid deposition fields (see for example Dore et al, 2015; Fagerli et al., 2003; NEGTAP, 2001; RoTAP, 2012). However, any comparison of model results with observations is never perfect. Inevitably, there is likely to be good agreement for some sites or species and not with others. There are inadequacies and simplifications in the model together with site dependent factors influencing the observations. Here, the view is taken that it is difficult to find a set of model input parameters that uniquely fit the available observations. There may be a number of sets of parameters, or combinations of parameters that are 'acceptably' consistent with the available observations. This is known as equifinality (Beven, 2006) and results from the difficulty of deciding between competing parameter sets and models, given the limitation of the observations. Equifinality implies uncertainty and is the basis for our exploration of uncertainty within HARM. We have approached this by adopting the Generalised Likelihood Uncertainty Estimation (GLUE) framework.

In a previous study using HARM, Page et al. (2008) identified a subset of 11,699 HARM model runs that 'adequately' represented observed acid deposition data, allowing the production of deposition uncertainty distributions across the UK. This subset of 'acceptable' model parameter sets has been used in this study to provide distributions of deposition for 2005 and 2020. Details of the parameter set 'acceptance' criteria and the Monte Carlo parameter set sampling procedure are given in Page et al. (2008).

2.2 Critical loads for soils

Critical loads for soils were defined and estimated using the steady state mass balance method for GB (Hornung et al., 1995). CLs were assigned using the dominant soil type at a spatial scale of 1 km x 1 km using the Skokloster categories Class 1 to Class 5 and their distribution across Great Britain (GB) is shown in Figure 1. Class 1 soils have the lowest buffering capacity (most sensitive) and were assigned CLs in the range 0 – 0.2 keq ha^{-1} yr^{-1}. Class 5 soils have the highest buffering capacity and were assigned CLs greater than 4.0 keq ha^{-1} yr^{-1}. Soils in Classes 2, 3 and 4 have intermediate levels of buffering capacity and had their range boundaries set at 0.5, 1.0 and 2.0 keq ha^{-1} yr^{-1}. Given the difference in spatial scale between the CL data (1 km x 1 km) and the HARM deposition data (10 km x 10 km), the CL data were aggregated up to the scale of the HARM data, providing the total area for each Skokloster soil class within each 10 x 10 km grid cell. Aggregating up the CLs in this way does not change the underlying sensitivity, but masks the spatial distribution and location of the most sensitive elements within each square. This spatial distribution is only important if there are strong gradients in deposition within a particular grid square or the assessment of damage is required for a particular location. At the 10 km x 10 km scale such gradients were not significant and hence the aggregation process led to no significant loss of accuracy or bias in the CL exceedance.

In total, there were 1467 10 km x 10 km grid squares representing England, 258 for Wales and 1047 for Scotland. No corresponding CL data were available for Northern Ireland and so this country was given no further consideration in this analysis. Here, the effects of incorporating uncertainties associated with the Skokloster CL classifications into the calculation of CL exceedances has been studied for the 2772 grid squares covering GB, given the uncertain deposition estimates described above.

Uncertainties in the estimation of CLs were first addressed by Zak et al. (1997) who applied the GLUE approach to the PROFILE model, a steady state geochemical model that is widely used within the CL community. Heywood et al. (2006b) used coniferous woodland as an example and showed that uncertainties in GB CLs varied between 14 – 29%. In further work, Heywood et al. (2006c) reviewed uncertainties in CL assessments across Europe and established the need for a coordinated effort to characterise uncertainties in CLs. Skeffington et al. (2007) used Monte Carlo methods to obtain the output distributions of various CL parameters, having quantified the uncertainties in the input parameters to the CL models. They showed that estimates of the uncertainties in the CLs for acidity exhibited coefficients of variation which lay between 25 and 61%, across a range of catchments. On the basis of the uncertainties estimated by Heywood et al. (2006b) and Skeffington et al. (2007), we take the view that the uncertainties in actual CLs are likely to be smaller, or at most comparable to, the ranges in the Skokloster classes outlined above.

The uncertainty in the CLs within each 10 km x 10 km grid square was addressed by assigning the CL a probability distribution that was evenly distributed within the particular CL range, that is to say, a 'top hat' function was assumed, as shown in Figure 2. As there was no HARM model estimated CL exceedance of the least sensitive (Class 5) soils in either 2005 or 2020, they are not discussed in this paper.

2.3 Estimating critical loads exceedances and their uncertainties

The methodology employed in the estimation of the uncertain CL exceedances for soils is illustrated in Figure 2. It consisted of a loop over the 2772 GB grid cells. Within this loop, the 11,699 acceptable

HARM estimates of total acid deposition for each 10km grid cell were overlaid onto the CL ranges for each soil class to estimate CL exceedances, as follows:

CL exceedance (keq ha^{-1} yr^{-1}) = acid deposition load (in keq ha^{-1} yr^{-1}) - CL (in keq ha^{-1} yr^{-1}).

The accumulated exceedance (AE) of the CLs in a given grid square was calculated using:

Accumulated Exceedance (keq yr^{-1}) = CL exceedance x area exceeded

and summing this over all the soil classes in a given grid square. This calculation was repeated for each of the soil classes and each of the 10 km x 10 km grid squares.

This methodology was then repeated using the 11,699 HARM deposition estimates for the 2020 emission scenario. For each soil class and grid square, the differences in AE (2005 – 2020) were calculated: these differences were calculated by pairing up the 11,699 HARM estimates for 2005 and 2020 and not drawing them at random from the sets of model runs. The differences in AE were then ranked in order and the 5th-, 25th-, 50th-, 75th- and 95th-percentiles were determined for the distributions of the 11,699 'acceptable' results.

3. Estimating 2005 – 2020 differences in critical load exceedance in the Peak District

To illustrate the application of the methodology in Figure 2, attention is turned to a single 10 km x 10 km grid square located in the Peak District National Park, in northern England (see inset Figure 1). Class 1 soils occupied 25% of the surface area of this grid square, Class 2 14%, Class 3 22% and Class 4 25%. Total HARM acid deposition declined from $1.29^{+0.59}_{-0.40}$ keq ha^{-1} yr^{-1} (where the quoted uncertainty range is the 5% - 95% range, equivalent to the 2 – σ confidence interval) in 2005 to $0.93^{+0.39}_{-0.29}$ keq ha^{-1} yr^{-1} in 2020, giving a reduction in acid deposition of $0.36^{+0.30}_{-0.11}$ keq ha^{-1} yr^{-1}.

The probability distribution of the HARM model estimates of the difference in AE per class is illustrated as a box-and-whisker plot in Figure 3. Looking first at the Class 1 (most sensitive) soils, all 11,699 model runs for both 2005 and 2020 gave deposition estimates that exceeded the CL for Class

1 soils. The 2005 – 2020 difference in AE for Class 1 soils was found to be 895^{+493}_{-290} keq yr^{-1}. On this basis, the 5% - 95% confidence interval was narrow enough not to encompass zero and it could be concluded that the difference in AE was statistically significantly different from zero, despite the uncertainties in the deposition and CLs. However, in Figure 3, it can be seen that the 2 – σ confidence interval was not exactly symmetrical about the 50-percentile value. This lack of symmetry implies a degree of skewness in the distribution of the differences in the AEs. Statements about statistical significance based on the assumption of a normal distribution may not be reliable if there is a high degree of skew. However, on a cautionary basis, if the range between the 50-percentile and the upper confidence limit was applied at the lower confidence interval, then the 5% - 95% range would still not encompass zero. It was thus concluded that the difference in AE was likely to be robust, despite the apparent skewness in its probability distribution and the uncertainties in the deposition and CLs.

The deposition loads exceeded the CLs for Class 2 soils in all HARM model runs in both 2005 and 2020. The AE for Class 2 soils was 1297^{+600}_{-442} keq yr^{-1} in 2005 and 795^{+500}_{-300} keq yr^{-1} in 2020, with a difference in AE of 501^{+276}_{-162} keq yr^{-1}. Since the 2 – σ confidence interval did not encompass zero, it was concluded that this difference was statistically significant, taking into account the apparent skewness in its probability distribution. The situation was much the same for Class 3 soils, where the 2005 – 2020 difference in AE was found to be 763^{+458}_{-394} keq yr^{-1}, see Figure 3, and again this difference was considered to be significantly different from zero.

Looking at the least sensitive Class 4 soils, all 11,699 model runs gave deposition estimates that exceeded the CL in 2005, but 75% of the model runs met critical loads in 2020. The 2005 – 2020 difference in AE was found to be 84^{+511}_{-84} keq yr^{-1}. The skewness in the distribution for the Class 4 soils is clearly apparent in Figure 3. Uncertainties were so large for the Class 4 soils that they encompassed zero and so it was unlikely that they could be considered significant because of the combined uncertainties in the deposition and CLs. We therefore have the situation where in one

10km grid square, the most sensitive soils show a large and statistically significant reduction in AE whereas the least sensitive soils show a small reduction, which is not significant. This contradicts our conventional notion of environmental protection that if you protect the most sensitive elements in the environment from damage, then you automatically protect the least sensitive. However, because CLs were actually met for Class 4 soils in three cases out of four, the small difference in AE and its lack of statistical significance would not be relevant in policy terms.

4. Estimating 2005 – 2020 differences in critical loads exceedance across GB

The methodology illustrated in Figure 2 was then followed for each of the 2772 10 km x 10 km grid squares across GB. We found that the differences in AE between 2005 and 2020 for all soil classes (1 – 4) showed that the reductions in emissions in our initial scenario reduced CL exceedances throughout GB. This implies that non-linearities in the relationship between acid deposition and CL exceedance were unimportant on the GB scale. This is a reflection of the illustrative emission reduction scenario chosen, where there was no reduction in the emissions of NH_3 across the UK and very limited (4%) reduction across the rest of the EMEP area, hence, non-linearities in relation to the response of S and oxidised N to changes in the emission of NH_x were minimised.

The 2005 – 2020 difference in total AE for Class 1 soils was $354{,}000\ ^{+145{,}000}_{-104{,}000}$ keq yr^{-1} (see Table 1) The probability distribution of the AE differences is shown as a box-and-whisker plot in Figure 4 and a 2 – σ confidence range did not encompass zero. Despite the uncertainties in the deposition loads and CLs, this difference in AE was statistically significant. The spatial distribution in the 50-percentile reductions in AE for the individual grid squares is shown in Figure 5a. The greatest reductions were found in southern England, Wales, East Anglia, northern England and in a few scattered locations in south west Scotland and in the highlands and islands. The 2 – σ ranges in the differences in AE for the individual grid squares were not evenly distributed about their 50-percentile values. The dispersion in the AEs about their 50-percentiles showed evidence of skewness, with shorter tails to low values and longer tails to high values (Figure 4). However, as with the Peak District grid square, this dispersion differed only slightly from that shown by a 'normal' distribution. Consequently, a null hypothesis that the AE reductions were due to chance could be rejected with a high level of confidence. On this basis, it was concluded that the reductions in the AEs for Class 1 soils were all highly significant at the 99.99% level, despite the large uncertainties in the deposition loads and CLs. Although the changes for this soil class were small (Figure 4) they are likely to be important for these most acid sensitive environments. There were a small number of grid squares, on the fringes of GB, where it was difficult to make any robust statement about the policy significance of any reduction in AE because of severe skewness.

The difference in Total AE for Class 2 soils across GB was 1,275,000 $^{+460,000}_{-375,000}$ keq yr^{-1}, see Table 1 and Figure 4, between 3 – 4 times higher than for Class 1 soils. Again, the 2 – σ confidence range did not encompass zero and so this difference was highly statistically significant. Although CL exceedances were generally higher for Class 1 soils, the areas assigned to Class 2 soils were much larger and so the total AE difference across GB was substantially higher for the latter. Figure 5b shows the spatial distribution of the 50-percentile AE differences for Class 2 soils for each grid square. The greatest reductions in AE were found in Wales, Cumbria, south west Scotland and across the Scottish Highlands. Although the distributions in the AE differences were skewed, the degree of skewness was considerably less than for Class 1 soils (Figure 4). It was concluded that the reductions in the AEs for Class 2 soils were all highly significant at the 99.99% level, despite the large uncertainties in the deposition and CLs. Skewness was a real problem in less than 3% of grid squares, the bulk of these in the Outer Hebrides. It is difficult to make any robust statement about the environmental significance of the AE reduction in these locations.

The difference in total AE across GB for Class 3 soils was 1,010,000 $^{+780,000}_{-565,000}$ keq yr^{-1}, see Table 1 and Figure 4. This AE difference was somewhat smaller than for Class 2 soils despite their substantially larger areal coverage because of their lower CL exceedances. Although the 2 – σ confidence interval did not encompass zero, there was noticeable skewness in the distribution of AE differences. As discussed above, statements about significance may not be reliable if there is a large amount of skewness. However, as with the Peak District grid square, if the 50-percentile – 95-percentile range was applied at the lower confidence interval, then the adjusted 5-percentile – 95-percentile range would still not encompass zero. It was concluded that the difference in total AE was likely to be robust, despite the uncertainties in the deposition and CLs. Figure 5c shows the spatial distribution of the 50-percentile differences for the individual grid squares containing Class 3 soils. The largest reductions were found throughout southern and south west England, south Wales and a band from the west Midlands and into north west England. In all these regions, the reductions were likely to be highly significant. However in the regions where the reductions were much smaller and close to zero, skewness was again a real, issue. In ~ 25% of the grid squares, it was considered likely that the reductions in AE were not significant. This resulted from the situation where CLs and deposition loads were comparable in magnitude so the combination of uncertainties has become overwhelming in the estimation of these small AEs.

The difference in total AE across GB for Class 4 soils was found to be 42,000 $^{+275,000}_{-41,000}$ keq yr^{-1}, see Table 1 and Figure 4. The spatial distribution of the 50-percentile differences for the individual grid squares containing Class 4 soils is shown in Figure 5d. The difference in AE is small and highly uncertain (the 2-σ confidence range encompasses zero) compared with the above same values for Class 1 – 3 soils. Deposition and CLs were closely comparable in magnitude and so the uncertainties in these quantities have been magnified in the estimation of AE differences to the extent that AE and its differences have become unreliable indicators of ecosystem status for Class 4 soils. Given the relative insensitivity of this class of soils to acidification it is, however, quite feasible that the 2020 scenario would deliver ecosystem protection.

5. Discussion and Conclusions

In the Introduction, we posed two policy related questions: The first question was if the current models and the current CL approaches are too uncertain to identify whether proposed emissions reductions will deliver discernible environmental improvement; the second question concerned the impact of the change in the optimisation target from CL exceedance to accumulated exceedance. We have applied the GLUE methodology to address the uncertainties in deposition models and in the CLs. We have then developed a realistic hypothetical scenario for 2020 and quantified the uncertainties in the estimates of the differences in AE between 2005 and 2020. The 2-σ confidence limits for the AE difference for Class 1 – 3 soils in the vast majority of GB locations do not encompass zero (see Figure 4) and so are likely to be statistically significant. In relation to question one, we can therefore say with some confidence that reductions in emissions of the order of 35% will lead to reductions in AE which are not 'lost in the noise' in the deposition and CL modelling. These findings are consistent with those of other studies for the UK (Helliwell et al., 2014; Majeko et al., 2009; Oxley et al., 2013;) using a range of modelling approaches. It is notable, however, that only the Helliwell et al. study (using the MAGIC model) attempted to include uncertainty in their assessment, primarily in relation to model inputs (parametric uncertainty). Far from being too uncertain for policy use, we have been able to make a first attempt at quantifying uncertainties in both deposition and CL at the GB scale and to demonstrate that the uncertainties are small enough that they can be employed to develop robust policy assessments. To follow on from Uusitalo et al. (2015, see Introduction) we can use this approach to give policy makers a more realistic picture of possible outcomes in tackling this particular 'wicked problem'.

The second question concerned the impact of the change in environmental target from simple CL exceedance (or not), to an index of success represented by AE. Using the standard CL approach, with a single value applied to a deposition grid cell, the degree of protection was assessed only on a true or false basis (see Introduction). If the outcome of running a future emissions scenario was false (ie CL was still exceeded), policy makers were left with the impression that the proposed emissions reductions would fail to deliver environmental protection. In contrast, using the AE index gives a broader measure of better or worse relative to the starting situation, even if CL are not met completely. In our 2020 scenario, based on our 11,699 model runs, CLs for Class 4 soils would be met 98% of the time. For Class 3 soils this declined to 67%, for Class 2 soils to 27% and for Class 1 soils (most sensitive) to slightly less than 1% (fewer than 116 runs of the 11,699). Only on the most extreme deposition and CL uncertainty outcomes would Class 1 and 2 soils be protected. This suggests that emissions reductions in line with current commitments would do little to protect the most acid sensitive environments across GB (see Table 1). A simple estimate of the magnitude of emission reduction needed to provide full protection (based on extrapolation from the 2020 results) indicated that an emission reduction of around 45% would be needed to protect Class 4 soils completely (compared with 35% in our 2020 scenario) and of around 85% for Class 3 soils. Only very extreme (and probably impractical) reductions would offer protection to the most sensitive soils (Class 1). The change of optimisation target from meeting CL to the use of AE has, however, allowed us to make progress in terms of policy assessment for the most sensitive soils in the face of uncertainties in deposition models and the CLs themselves.

As the science in deposition modelling and CL assessments develops, there should be a narrowing (bounding) of uncertainties (see Introduction) and this should lead to a narrowing of the uncertainties in the emission reductions required to meet critical loads for Class 1 soils. There are reasons to suppose that some deposition estimates for GB have been overestimated (Dore et al., 2015; see Hall and Smith 2015 for a specific example) and so our conclusions may well have underestimated the likely improvement in environmental protection afforded by our initial hypothetical emission scenario. It could be, however, that current emissions reduction targets will never be able to protect the most acid sensitive environments and that the recovery of both aquatic and terrestrial ecosystems could take decades, in spite of the marked decrease in exceedance since the peak in the 1970s and 1980s (De Wit et al., 2015).

The importance of both considering and communicating uncertainty has come to the fore recently because of the debate around this issue in relation to anthropogenic climate change. The idea that a quantitative approach to uncertainty should be incorporated into environmental policy making has, however, been around for more than 20 years (see Frey, 1992 in relation to the US EPA). As Cooke (2015) observes 'There are formidable pitfalls when reasoning under uncertainty, into which both the scientific community and the general population repeatedly fall' (p. 8), but there is no doubt that handling uncertainty in its various forms is now a key part of developing environmental policy in a variety of domains, as was suggested by Schneider and Kuntz-Duriseti (2002). We have set out one approach to achieving this, focusing on the implications of taking uncertainty into account in controlling emissions of acidifying pollutants. It should certainly play a part in developing strategies for policy initiatives such as the latest iteration of the Convention on Long-range Transboundary Air Pollution (Gothenburg Protocol, see Introduction) as it attempts to provide the scientific basis and an effects based approach to addressing a widening range of atmospheric pollutant issues and their interactions with climate change and biodiversity (UNECE, 2016). The point of this study was to show how uncertainties could be handled rather than to make a formal assessment of acid deposition policies, but it is evident that in this case, as in others, uncertainty cannot be used as a reason to limit action (Drouet et al., 2015).

Acknowledgements

Original development of HARM was supported by UK Department of Environment, Food and Rural Affairs (DEFRA) and the UK Environment Agency.

References

Amann, M., Bertok, I., Borken-Kleefeld, J.,Cofala, J., Heyes, C., Höglund-Isaksson, L.,Klimont, Z., Rafaj, P., Schöpp, W. and Wagner, F. 2012. Environmental Improvements of the 2012 Revision of the Gothenburg Protocol. CIAM Report 1/2012. IIASA.

Battarbee, R.W., Shilland, E.M., Kernan, M., Monteith, D.T. and Curtis, C.J. 2014. Recovery of acidified surface waters from acidification in the United Kingdom after twenty years of chemical and biological monitoring (1988 – 2008) Ecological Indicators 37, 267-273.

Beven , K. 2006. A manifesto for the equifinality thesis. *Journal of Hydrology* 320, 18-36.

Bull, K.R. 1992 An introduction to critical loads. *Environmental Pollution* 77, 173-176.

Cooke, R.M. 2015. Messaging climate change uncertainty. *Nature Climate Change* 5, 8-10. doi:10.1038/nclimate2466

DEFRA 2015. Emissions of air quality pollutants 1990 – 2013. https://uk-air.defra.gov.uk/assets/documents/reports/cat07/1511261127_AQPI_Summary_1990-2013_Issue_v1.1.pdf

De Wit, H., Hettelingh, J.P. and Harmens, H. 2015. Trends in ecosystem health and responses to long-range transported atmospheric pollutants. ICP Waters report 125/2015.

Dore, A.J., Carlslaw, D.C., Braban, C., Cain, M., Chemel,. C., Conolly, C., Derwent, R.G., Griffiths, S.J., Hall, J., Hayman, G., Lawrence, S., Metcalfe, S.E., Redington, A., Simpson, D., Sutton, M.A., Sutton, P., Tang, Y.S. Vieno, M., Werner, M.and Whyatt, J.D. 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.

Drouet, L., Bosetti, V. and Tavoni, M. 2015. Selection of climate policies under the uncertainties in the Fifth Assessment Report of the IPCC. *Nature Climate Change* 5, 937-940. doi:10.1038/nclimate2721

European Environment Agency 2015.European Union emission inventory report 1990 – 2013 under the UNECE Convention on Long-range Transboundary Air Pollution (LRTAP). EEA Technical Report No. 8/2015.

Fagerli, H., Simpson, D., Aas, W., 2003. Model performance for sulphur and nitrogen compounds for the period 1980 to 2000. In: Tarrason L. (Ed.). Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe. EMEP Status Report 1/2003, Part II Unified EMEP Model Performance. The Norwegian Meteorological Institute, Oslo, Norway.

Frey, H.C. 1992. Quantitative analysis of uncertainty and variability in environmental policy making. Report for the AAAS/EPA.

Hall, J. and Smith, R. 2015. Trends in critical load exceedances in the UK. Report to DEFRA under contract AQ0826.

Helliwell, R.C., Aherne, J., MacDougall, G., Nisbet, T.R., Lawson, D., Cosby, B.J. and Evans, C.D. 2014. Past acidification and recovery of surface waters, soils and ecology in the United Kingdom: Prospects for the future under current deposition and land use protocols. *Ecological Indicators* 37, 381-395.

Hettelingh, J-P., Posch, M., De Smet, P.A.M and Downing, R.J. 1995.The use of critical loads in emission reduction agreements in Europe. *Water, Air and Soil Pollution* 85, 2381-2388.

Hettelingh, J-P., Posch, M., Velders, G.J.M. et al., 2013a. Assessing interim objectives for acidification, eutrophication and ground-level ozone of the EU National Emissions Ceilings Directive with 2011 and 2012 knowledge. *Atmospheric Environment* 75, 129-140. Doi:10.106/j.atmosenv.2013.03.060.

Hettelingh, J-P., Posch, M., Sllotweg, J. and Le Gall, A-C., 2013b. Assessing the effects of the revised Gothenburg Protocol. In: *Modelling and Mapping of Atmospherically-induced Ecosystem Impacts in Europe*, Posch, M., Slootweg, J. and Hettelngh, J-P. (eds), CCE Status Report 2012, chapter 1, pp. 13-20.

Heywood, E., Whyatt, J.D., Hall, J., Wadsworth, R. and Page, T.2006a. Presentation of the influence of deposition uncertainties on acidity critical load exceedances across Wales. *Environmental Science and Policy* 9, 32-45.

Heywood, L., Hall, J., Smith, R., 2006b. Uncertainty in mass balance critical loads and exceedance: Application to a UK national data set. *Atmospheric Environment* 40, 6146-6153.

Heywood, E., Hall, J., Reynolds, B., 2006c. A review of uncertainties in the inputs to critical loads of acidity and nutrient nitrogen for woodland habitats. *Environmental Science and Policy* 9, 78-88.

Holmberg, M.,Vuorenmaa, J., Posch, M., Fosius, M., Lundin. L., Kleemola, S., Augustaitis, A., Beudert, B., de Wit, H.A., Dirnbock, T., Evans, C.D., Frey, J., Grandin, U., Indriksone, I., Kram, P., Po,pei, E., Schulte-Bisping, H., Srybny, A. and Vana, M. 2013. Relationship between critical load exceedances and empirical impact indicators at Integrated Monitoring sites across Europe. *Ecological Indicators* 24, 256-265. 2013.

Hornung, M., Bull, K.R., Cresser, M., Hall, J., Langan, S.J., Loveland, P. and Smith, C., 1995. An empirical map of critical loads of acidity for soils in Great Britain. Environmental Pollution 90, 301-310.

Kernan, M., Battarbee, R.W., Curtis, C.J. et al. (Eds.). 2010 Recovery of lakes and streams in the UK from the effects of acid rain. UK Acid Waters Monitoring Network 20 Year Interpretative Report. ECRC, London.

Matejko, M., Dore, A.J., Hall, J., Dore, C.J., Blas, 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 and Policy* 12, 882-896.

Metcalfe, S.E., Whyatt, J.D., Nicholson, J.P.G., Derwent, R.G. and Heywood, E. 2005. Issues in model validation: assessing the performance of a regional-scale acid deposition model using measured and modelled data. *Atmospheric Environment* 39, 587-598.

NEGTAP, 2001. Transboundary Air Pollution. Acidification, Eutrophication and Ground-Level Ozone in the UK.

Oxley T., Dore, A.J., ApSimon, H., Hall, J. and Kryza, M. 2013. Modelling future impacts of air pollution using the multi-scale UK Integrated Assessment Model (UKIAM). *Environment International* 61, 17-35.

Page, T., Whyatt, J.D., Metcalfe, S.E., Derwent, R.G., Curtis, C., 2008. Assessment of uncertainties in a long range atmospheric transport model: Methodology, application and implications in a UK context. *Environmental Pollution* 156, 997-1006.

Posch, M., Slootweg, J. and Hettelingh, J-P. (eds.) 2012. Modelling and mapping of atmospherically-induced ecosystem impacts in Europe. CCE Status Report 2012.

Refsgaard, J.C., van der Sluijs, J.P., Lajer Højberg, A. and Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process – A framework and guidance. Environmental Modelling and Software 22, 1543-1556.

Reis, S., Grennfelt, P., Klimont, Z., Amann, M., ApSimon, H., Hettelingh, J-P., Holland, M., LeGall, A-C., Maas, R., Posch, M., Spranger, T., Sutton, M. and Williams, M. 2012. From acid rain to climate change. Science 338, 1153-1154.

RoTAP, 2012. Review of Transboundary Air Pollution: Acidification, Eutrophication, Ground Level Ozone and Heavy Metals in the UK. Contract Report to DEFRA.CEH.

Schneider, S.H. and Kuntz-Duriseti, K., 2002. Uncertainty and climate change policy. In: Schneider, S.H., Rosencranz, A. and Niles, J.O. (eds.) Climate Change Policy: A Survey. Island Press, Washington DC, pp. 53-87.

Slootweg, J., Posch, M. and Hettelingh, J-P.(eds) 2015. Modelling and mapping the impacts of atmospheric deposition of nitrogen and sulphur: CCE Status Report 2015, Coordination Centre for Effects.

Skeffington, R.A., Whitehead, P.G., Heywood, E., Hall, J.R., Wadsworth, R.A., Reynolds, B., 2007. Estimating uncertainty in terrestrial critical loads and their exceedances at four sites in the UK. Sci. Total Environ. 382, 199-213.

UNECE 2016 Decision 2010/18 Long-term strategy for the Convention on Long-range Transboundary Air Pollution and Action Plan for its Implementation. ECE/EB.AIR/016/Add.1

Uusitalo, L., Lehikoinene, A., Helle, I and Myrberg, K., 2015. An overview of methods to evaluate uncertainty of deterministic models in decisions support. Environmental Modelling and Software 63, 24-31.

Whyatt, J.D, Metcalfe, S.E., Nicolson, J., Derwent, R.G., Page, T. and Stedman, S. 2007. Regional scale modelling of particulate matter in the UK, source attribution and assessment of uncertainties. *Atmospheric Environment* 41, 3315-3327.

Zak, S.K., Beven, K., Reynolds, B., 1997. Uncertainty in the estimation of critical loads: A practical methodology. *Water, Air, and Soil Pollution* 98, 297-316.

FIGURES

- Figure 1 – Single Column
- Figure 2 – Double Column (for legibility)
- Figure 3 – Single Column
- Figure 4 – Single Column
- Figure 5 – Double Column (4 maps)

Figure 1. Critical loads in keq ha^{-1} yr^{-1} for the dominant soil type at a spatial scale of 10 km x 10 km for Great Britain using the Skokloster categories Class 1 (most sensitive: in black) to Class 5 (least sensitive: in blue) estimated using the steady state mass balance method (Hornung et al., 1995). Inset shows detail for Peak District grid square.

Figure 2. A sketch illustrating the methodology adopted for the estimation of the frequency distributions of the differences in accumulated critical loads exceedance in a given 10km grid square between 2005 and 2020. The upper plots show the CL ranges for individual soil classes as coloured bars, a) Class 1, b) Class 2, c) Class 3, d) Class 4. The divisions within these bars indicate sampling within these ranges. The upper middle plots show accumulated exceedance for each individual soil class under the 2005 (in black) and 2020 (in blue) scenarios. The lower middle plots show the difference (reduction) in accumulated exceedance for each individual soil class between 2005 and 2020. The bottom plot (e) shows accumulated exceedance for all soil classes under the 2005 (black) and 2020 (blue) scenarios.

Figure 3. Box-and-whisker plots of the dispersion in the estimates of the reductions in accumulated exceedance between 2005 and 2020 for each soil class in the Peak District grid cell.

Figure 4. Box-and-whisker plots of the dispersion in the estimates of the reductions in

accumulated exceedance between 2005 and 2020 for each soil class across GB.

Figure 5. Spatial variations in the 50-percentile points of the distribution of the estimates of the reduction in accumulated CL exceedance between 2005 and 2020 for a) Class 1 soils, b) Class 2 soils, c) Class 3 soils and d) Class 4 soils.

TABLES

Table 1. Percentile points in the reduction in AE between 2005 and 2020 for each Skokloster soil class across GB in keq yr^{-1}.

Table 1.

Percentile	Class 1	Class 2	Class 3	Class 4	All classes
5%-ile	250,000	900,000	445,000	1,000	1,596,000
16%-ile	283,000	1,030,000	620,000	6,000	1,939,000
25%-ile	303,000	1,100,000	725,000	12,000	2,140,000
50%-ile	354,000	1,275,000	1,010,000	42,000	2,681,000
75%-ile	415,000	1,465,000	1,345,000	111,000	3,336,000
84%-ile	445,000	1,565,000	1,515,000	167,000	3,692,000
95%-ile	499,000	1,735,000	1,790,000	317,000	4,341,000

Figure 1

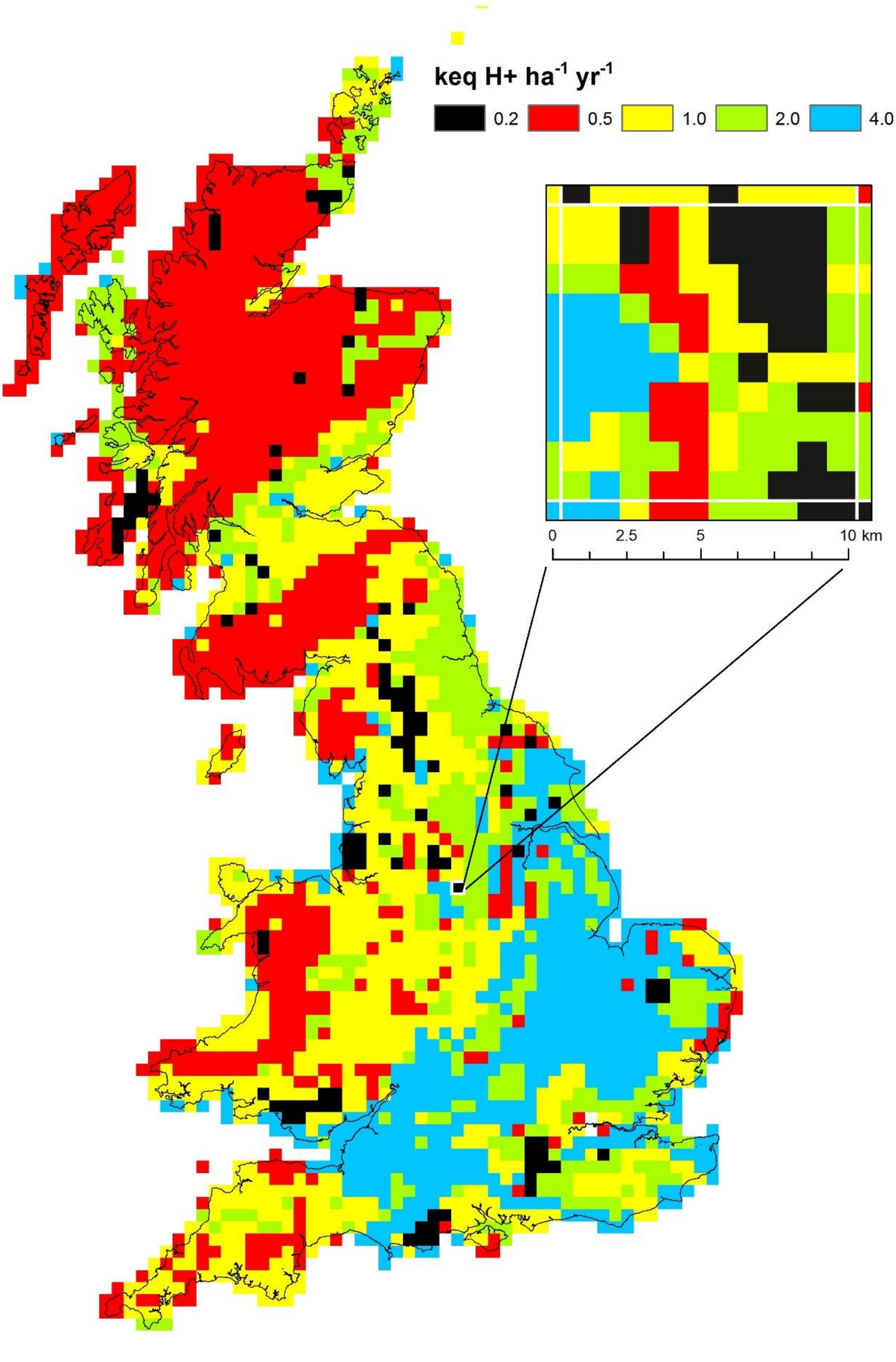
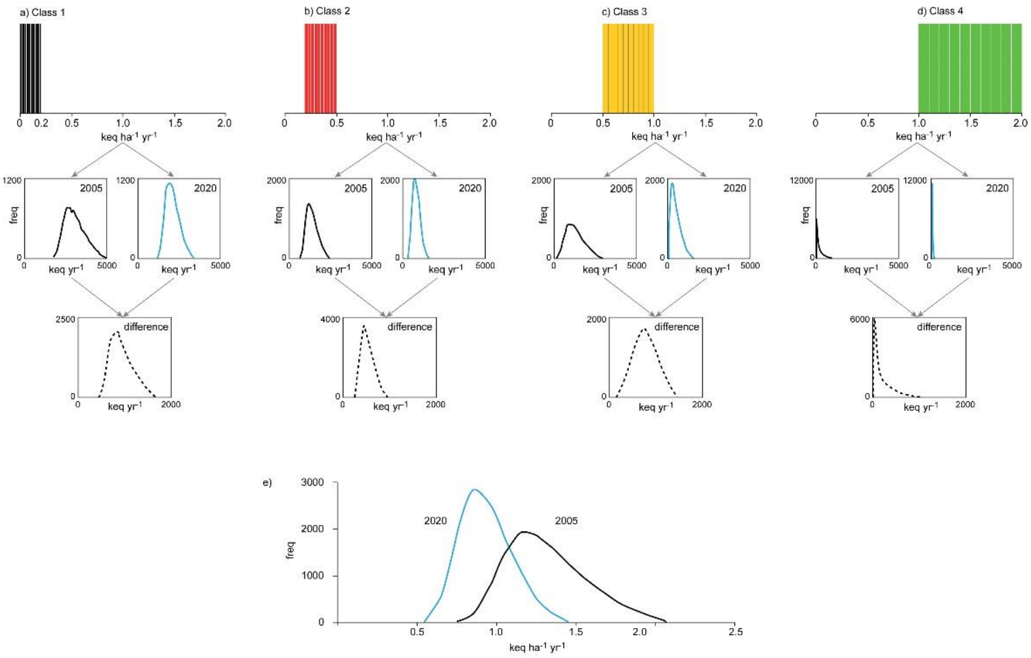
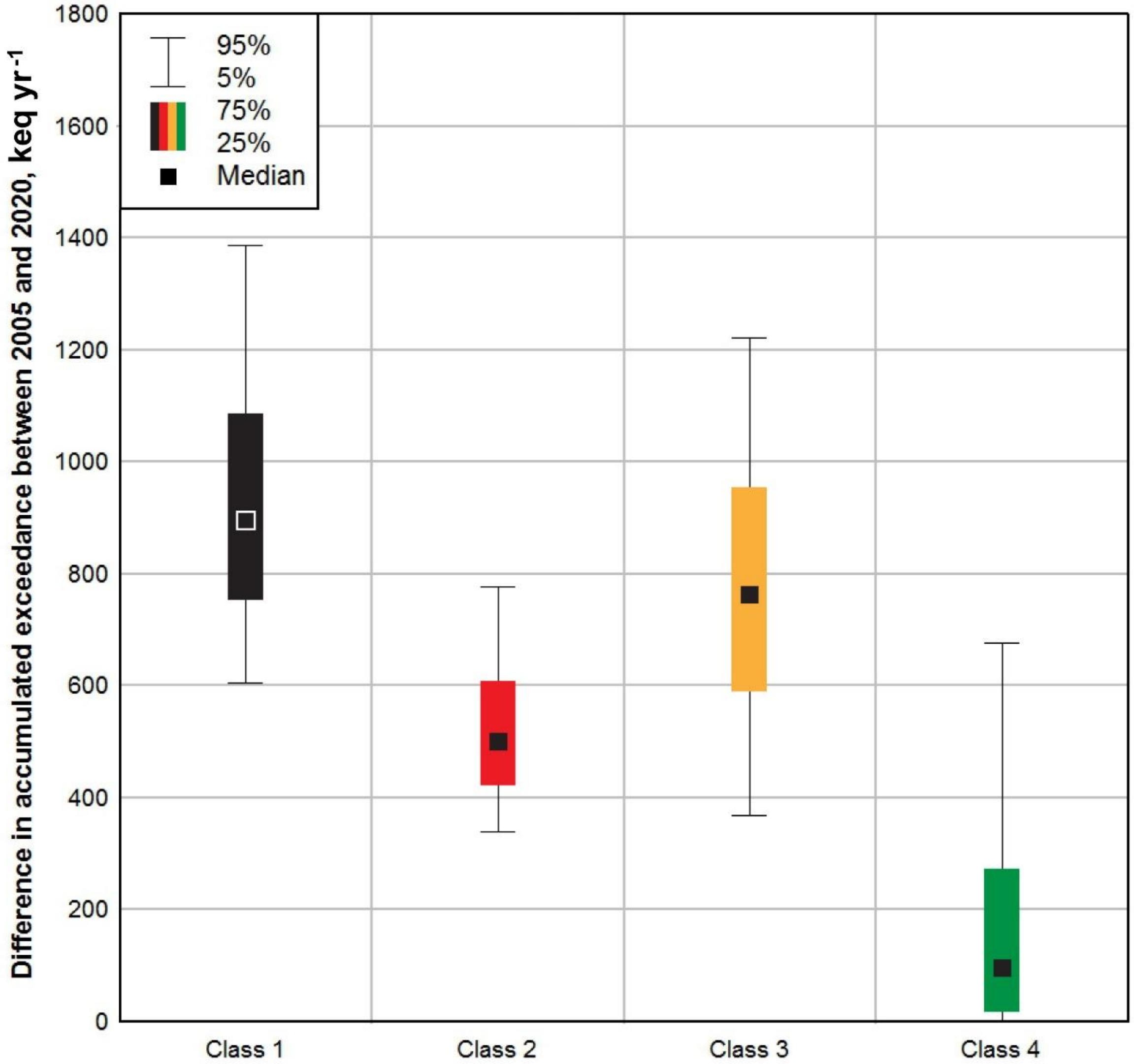


Figure 2



a) Class 1
b) Class 2
c) Class 3
d) Class 4
keq ha-1 yr-1
2005
2020
freq
keq yr-1
difference
e)

Figure 3



Difference in accumulated exceedance between 2005 and 2020, keq yr-1
95%
5%
75%
25%
Median
1800
1600
1400
1200
1000
800
600
400
200
0
Class 1
Class 2
Class 3
Class 4

Figure 4

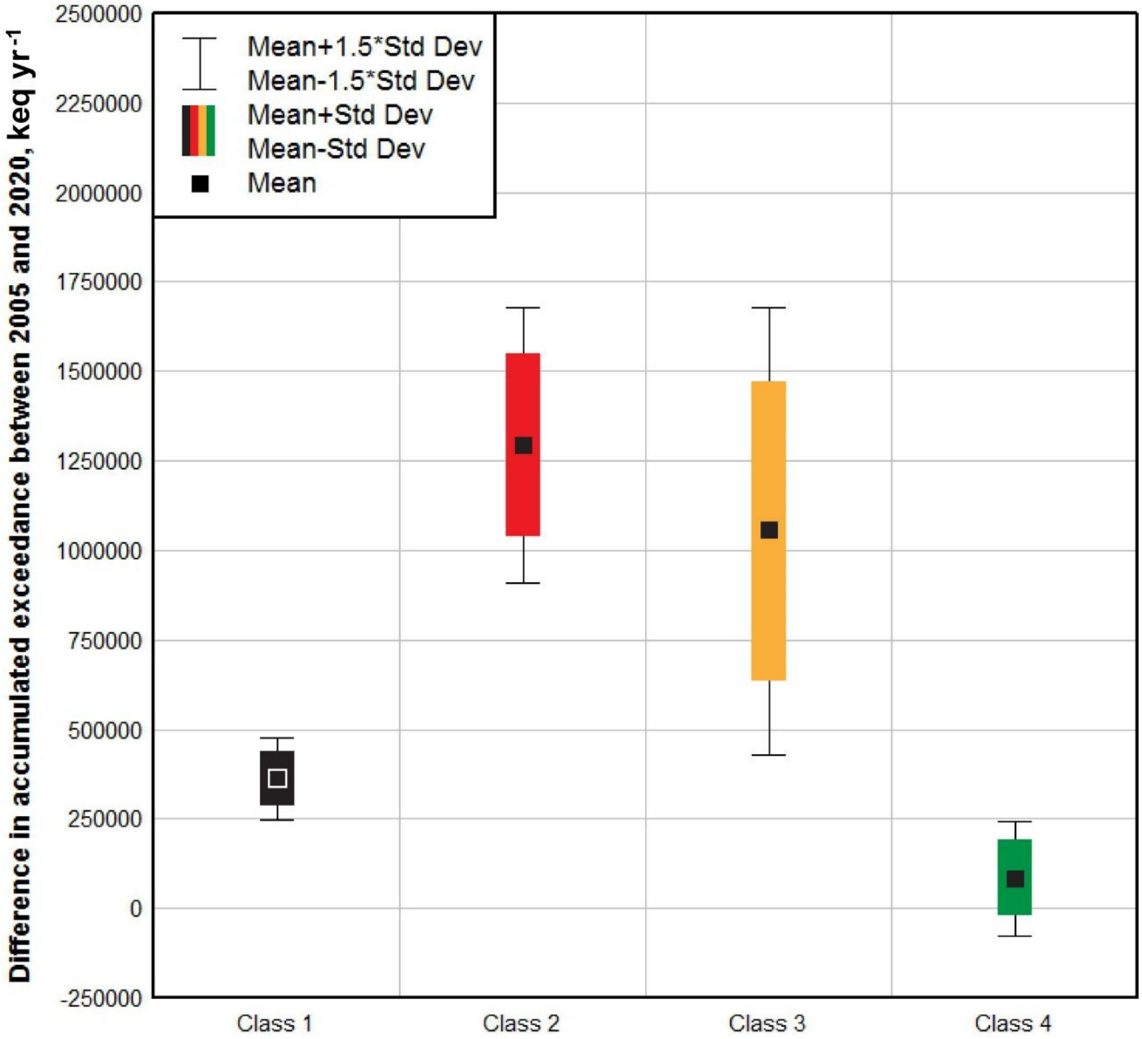
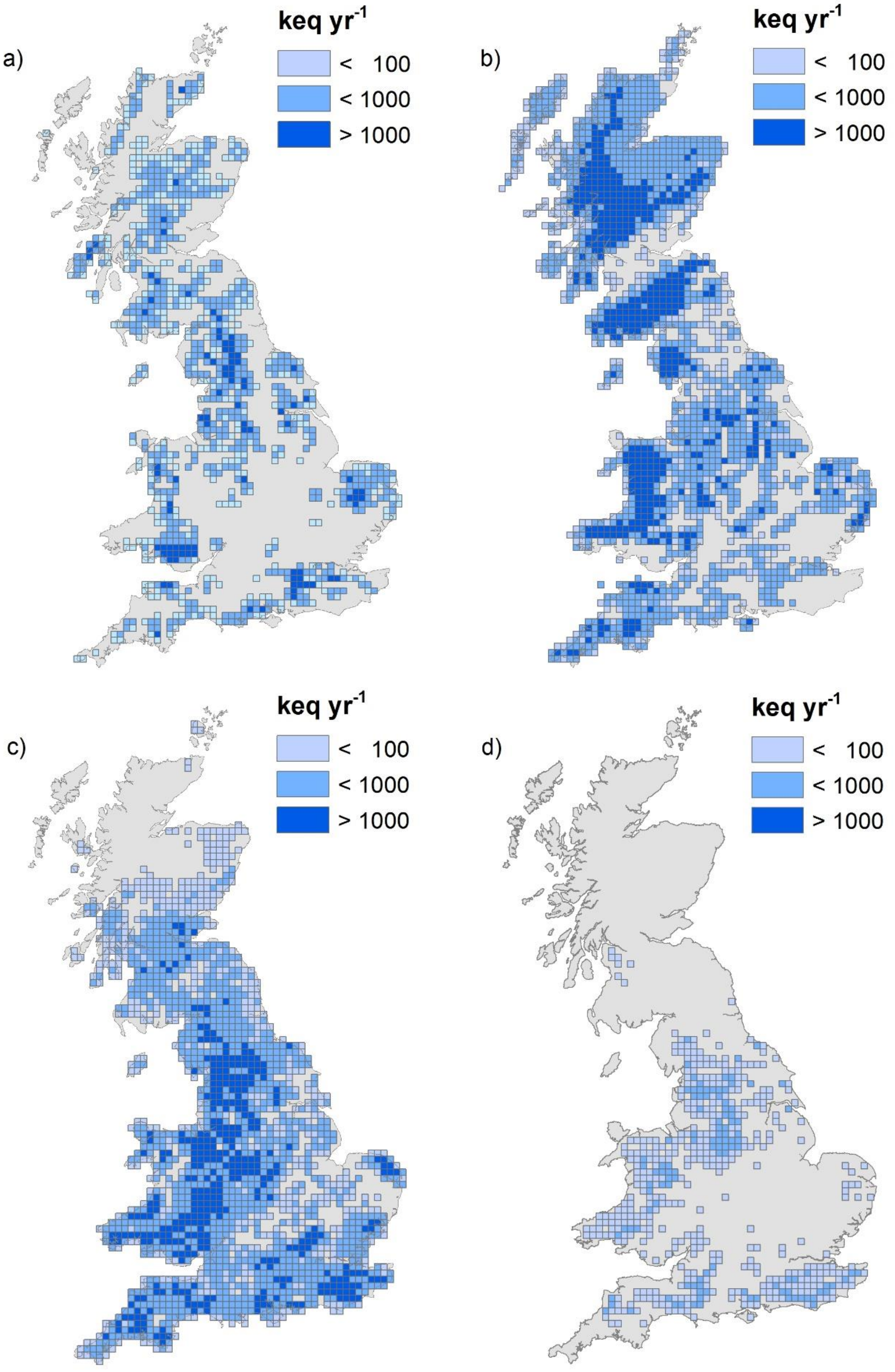

Mean+1.5*Std Dev
Mean-1.5*Std Dev
Mean+Std Dev
Mean-Std Dev
Mean
Difference in accumulated exceedance between 2005 and 2020, keq yr-1
2500000
2250000
2000000
1750000
1500000
1250000
1000000
750000
500000
250000
0
-250000
Class 1
Class 2
Class 3
Class 4

Figure 5


a)
keq yr^{-1}
< 100
< 1000
> 1000
b)
keq yr^{-1}
< 100
< 1000
> 1000
c)
keq yr^{-1}
< 100
< 1000
> 1000
d)
keq yr^{-1}
< 100
< 1000
> 1000

Vitae

Duncan Whyatt is a senior lecturer at Lancaster University. He is a geographer with over 25 years' experience of applying geospatial techniques in environmental research at local, national and regional scales. He uses GIS to visualise and analyse spatial data from different sources including pollution models. He has expertise in running a range of models to address different aspects of air pollution.

Sarah Metcalfe is Professor of Earth and Environmental Dynamics in the School of Geography at the University of Nottingham, UK. She has worked on modelling air pollution in the UK context for many years. She served on a number of scientific advisory groups for the UK government including the Review Group on Acid Rain, the Critical Loads Advisory Group and the National Expert Group on Transboundary Air Pollution and carried out research for the UK's devolved administrations and the Environment Agency.

Professor Richard (Dick) Derwent took a degree in 1968 and a PhD in 1971 from the University of Cambridge in physical chemistry. Dick Derwent has spent much of his research career studying air pollution. Initially, this carried out in the Air Pollution Division, Warren Spring Laboratory, then at the Harwell Laboratory and finally at the Meteorological Office, Bracknell. In 2003, he took early retirement and became a self-employed consultant on air pollution.

Trevor Page is a senior research associate at Lancaster University, UK. His interests are primarily in environmental systems modelling with a focus on hydrological and geochemical fluxes through catchments. Specifically, his work includes model uncertainty analyses coupled with evaluating the value of different types of data for improving model process-representation and model predictions. Much of his work has utilised Generalised Likelihood Uncertainty Estimation as a framework for these assessments.