Environmental Pollution 179 (2013) 285–293

Contents lists available at SciVerse ScienceDirect

Environmental Pollution

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

A systematic review of the effectiveness of liming to mitigate impacts of river acidification on fish and macro-invertebrates☆

Rebecca C. Mant[a,*], David L. Jones[a], Brian Reynolds[b], Steve J. Ormerod[c], Andrew S. Pullin[a]

[a] School of Environment, Natural Resources & Geography, Bangor University, Deiniol Road, Bangor LL57 2UW, UK
[b] Centre for Ecology & Hydrology, Environment Centre Wales, Deiniol Road, Bangor LL57 2UW, UK
[c] Catchment Research Group, Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK

ARTICLE INFO

Article history:
Received 16 August 2012
Received in revised form
19 January 2013
Accepted 11 April 2013

Keywords:
Acid rain
Alkalinity
Meta analysis
Salmonids
Invertebrates

ABSTRACT

The addition of calcium carbonate to catchments or watercourses – liming – has been used widely to mitigate freshwater acidification but the abatement of acidifying emissions has led to questions about its effectiveness and necessity. We conducted a systematic review and meta-analysis of the impact of liming streams and rivers on two key groups of river organisms: fish and invertebrates. On average, liming increased the abundance and richness of acid-sensitive invertebrates and increased overall fish abundance, but benefits were variable and not guaranteed in all rivers. Where B-A-C-I designs (before-after-control-impact) were used to reduce bias, there was evidence that liming decreased overall invertebrate abundance. This systematic review indicates that liming has the potential to mitigate the symptoms of acidification in some instances, but effects are mixed. Future studies should use robust designs to isolate recovery due to liming from decreasing acid deposition, and assess factors affecting liming outcomes.

1. Introduction

In the 1970's widespread environmental concern developed over the effects of acid deposition – widely known as "acid rain" – on base-poor streams, rivers and lakes (Menz and Seip, 2004). Sulphur and nitrogen oxides released into the atmosphere from industrial emissions decreased rainfall pH over affected areas and increased sulphate and nitrate concentrations in deposition. Where rocks and soils were base-poor, base-cation depletion and soil acidification followed, runoff pH declined, and concentrations of aluminium and other metals increased in soil and stream waters as explained by the well-known 'mobile anion' hypothesis (Reuss and Johnson, 1986). Surface-water acidification also changed many aspects of freshwater ecosystems, with altered invertebrate taxonomic composition and reduced fish populations, notably salmonids, among the best known effects (Moiseenko, 2005; Sandøy and Langåker, 2001; Schindler et al., 1985; Watt, 1987). At its peak, acid deposition was one of the most widespread pollution problems affecting rivers, and in base-poor locations such as Wales around half the stream length – some 12,000 km – were impacted (Firth et al., 1995).

Since the 1970s, industrial emissions have declined both in Europe and North America as a consequence of de-industrialisation and improved regulation leading to reduced concentrations of (non-marine) sulphate in runoff (Evans et al., 2001; Reynolds et al., 2004). However, emissions of nitrogen oxides have not decreased to the same extent (Fowler et al., 2007), and there are regions where acid deposition still exceeds soil neutralizing capacity (Matejko et al., 2009). Moreover, biological recovery in watercourses has been patchy or partial even in locations where mean runoff pH has increased (Ormerod and Durance, 2009). Currently, the best explanation for these circumstances is that episodic acidification still occurs during high discharge and is sufficient to offset biological recovery (Evans et al., 2008; Kowalik et al., 2007). At other locations, chronic acidification still remains a problem (Ormerod and Durance, 2009). Additionally, as industrialisation has shifted from Europe into South Asia, acid deposition has become an issue in other parts of the world (Kuylenstierna et al., 2001). In combination, these circumstances raise the possibility that liming – long advocated as a means of treating the symptoms of acidification (Clair and Hindar, 2005) – might be suggested more widely to protect waters where acidic deposition is a growing problem or to aid recovery where this is impaired.

* Corresponding author.
E-mail addresses: rebecca.mant@cantab.net, beccymant@gmail.com (R.C. Mant), d.jones@bangor.ac.uk (D.L. Jones), br@ceh.ac.uk (B. Reynolds), Ormerod@cardiff.ac.uk (S.J. Ormerod), a.s.pullin@bangor.ac.uk (A.S. Pullin).

0269-7491/$ – see front matter © 2013 The Authors. Published by Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.envpol.2013.04.019

Liming – the addition of calcium carbonate – is intended to raise the pH of rivers and/or lakes and occurs through a range of different methods. Limestone can be added directly in bulk into the river channel (termed point application in this review), applied continuously by mechanical dosers, applied directly to lakes within river catchments (lake liming) or distributed over river catchments (catchment liming). The latter is potentially effective in reducing the release of potentially toxic metal ions (e.g. Al^{3+}) from catchment soils (Clair and Hindar, 2005). Catchment liming can also be expected to have longer-term effects than individual direct applications although there are risks to the functioning and diversity of wetland systems that might be naturally acidic (Donnelly et al., 2003). With all liming methods, the most commonly used material is ground limestone gravel or powder, although dolomite, $CaMgCO_3$, is also used occasionally (Clair and Hindar, 2005). The dose applied can vary and is generally calculated by modelling the neutralizing requirements (Donnelly et al., 2003).

Liming has been implemented in North America and many European countries with some of the largest programs in Norway and Sweden (Clair and Hindar, 2005). The practice is still widespread in Europe despite reductions to some liming operations in Scandinavia as acid deposition has abated (Barlaup, 2004). For example, Sweden invested 3.8 billion SEK (approximately 0.4 billion Euros) on liming between 1983 and 2006 (Bostedt et al., 2010). Moreover, with the EU Water Framework Directive requiring member states to "protect, enhance and restore all bodies of surface water" to "good ecological status" (EU, 2000), there is the possibility that liming might be advocated through 'programmes of measures'.

For all the above reasons, it is timely to assemble the best evidence about liming effects to guide future applications and policy. While several long-term experiments have been carried out (e.g. Ormerod and Durance, 2009), there has previously been no systematic review appraising whether liming effectively restores fish and invertebrate populations in acidified rivers. In this paper, we provide such a systematic review, aiming to source, analyse and summarise the best available data. Specifically, we posed the question: "What is the impact of liming streams and rivers on the abundance and richness of fish and invertebrates?"

2. Materials and methods

2.1. Search for studies

A systematic review methodology was employed following standard guidelines (CEE, 2010). An a-priori protocol was completed and deposited in the Collaboration for Environmental Evidence Library (Mant et al., 2010). A systematic search for papers relevant to the question was then carried out using terms relevant to the focal ecosystem (i.e. streams/rivers), the biota (i.e. fish/invertebrates) and the intervention (liming). For each category, different variations of the terms were used in order to capture all relevant papers (Table 1). The search was conducted within ten databases including the Web of Knowledge and Scopus (Mant et al., 2011). Wherever the search engine allowed it, all search terms were used simultaneously. Terms within categories were linked with the Boolean operator 'OR', and terms between categories were linked with the Boolean operator 'AND' (Mant et al., 2011). No time, language or document type restrictions were applied. The use of English search terms could have biased the findings against papers in other languages. However, any such bias will have been reduced due to non-English language papers often still providing an English abstract and/or title.

Table 1
The search terms used to retrieve relevant papers, "*" denotes wildcard.

Population – ecological	Stream, river, catchment, brook, creek, burn, fluvial, source area, gravel.
Population–- biota	Fish* (includes fishes, fishery etc.), salmo*, trout, macroinvertebrate*, invert*, macrofauna, meiofauna, insect*, ephemeroptera, plecoptera, trichoptera, mollus*, crustacea*, microcrustacea*, bivalve*, gastropod, zooplankton, coleopteran, chironomid.
Intervention	Liming, lime*, chalk*, calcium carbonate, dolomite.

To find additional reports not retrieved by the database search, general web searches were conducted along with searches of the websites of relevant organisations including each of the Scandinavian, the US and the UK environment agencies (Mant et al., 2011). Bibliographies of material included were searched further for relevant references. Although review articles do not normally contain primary data, they were searched for any primary studies. No geographic restriction was applied to this review.

2.2. Study inclusion

Articles were assessed by their title, abstract then full text to identify those most relevant to the review using specific criteria. They were required to investigate change in abundance, density or richness of fish or invertebrates in any stream, river, or catchment where calcium carbonate (or dolomite) had been added to ameliorate the effects of anthropogenic acidification. Separating natural from anthropogenic effects on surface-water pH can be challenging, but relevant studies were those where liming was carried out to mitigate acidification that was assumed to be of human origin. Liming to mitigate acid mine drainage was not included. No particular method of liming was excluded. All primary studies that compared both limed and un-limed subjects were included, i.e. those which compared a limed river to the condition before liming or to a non-limed control (or both). At each stage, if there was insufficient information to exclude an article it was retained until the next stage. In order to assess and limit the effects of between-reviewer differences in determining relevance, two reviewers applied the inclusion criteria to 200 articles (over 20%) at the start of the abstract filtering stage. Analysis of agreement between the two was reasonable based on a kappa statistic of 0.6 (Edwards et al., 2002). Studies were excluded from the meta-analysis if relevant data could not be extracted due to incomplete reporting, lack of appropriate controls or multiple interventions being applied at the same time (6 articles). Additionally, for each river studied, the impact of liming was only recorded once for each outcome of interest, excluding 33 articles from analysis due to overlap in reported data.

Thirty-three relevant articles where included in the analysis, plus the main survey of the Norwegian environment agency (Direktoratet for naturforvaltning), and the main dataset of the Swedish environment agency (Naturvårdsverket). In total these 33 articles and 2 datasets covered 47 studies, 19 of which were rivers in the Norwegian survey and one of which was the main Swedish study that covered 18 limed rivers and 8 acid control rivers; details of all studies are given in Mant et al. (2011). Of the main 28 studies not in the Norwegian survey the majority (15) were on single rivers or streams and only three (all from Sweden) were on 10 or more rivers or streams. The studies included lake liming ($n = 4$ studies), catchment liming ($n = 6$), point applications into rivers ($n = 9$) and continuous dosing into rivers ($n = 9$). There were also 19 studies in which the liming method was unclear or multiple methods were used in different or the same river. In total there were 33 studies on fish and 27 studies on invertebrates, though in several both groups were assessed.

2.3. Data extraction and synthesis

All 47 studies included were appraised critically according to their study design and quality. Well-conducted studies of high quality have less potential for bias than their poorer counterparts. The presence of control and base-line data was recorded along with the level of replication, how the treatments and controls were allocated and the presence of confounding factors. Study outcomes were also recorded. Data on changes in invertebrate and fish abundance and species richness were extracted. Data on acid sensitive invertebrates, as defined by the study author, were also extracted. The impact of liming was calculated for each outcome (fish, invertebrates, acid sensitive invertebrates, richness and abundance) as the log ratio of limed to unlimed sites. The raw mean difference could not be used because units varied among studies. For fish population estimates units included density estimates of number of adult fish per 100 m^2 or number of fry and par per 100 m^2, fish biomass in kg/ha and fish biomass in total kg caught per year.

Meta-analyses were carried out on the extracted effect sizes using the *R* package 'metafor' (Viechtbauer, 2010). Random effects meta-analyses, weighting by inverse variance, were carried out using the DerSimonian–Laird estimator method. The weighted mean effects, confidence in the mean effect and prediction interval were calculated for each of the variables analysed. The confidence interval for the mean effect was the interval in which we were 95% confident that the mean effect occurred (i.e. the average effect across multiple studies). As several factors varied between studies, including physical, chemical and ecological characteristics of the rivers, not all liming operations will have produced the mean effect; the study-specific "true effect" varied between studies. Hence, the prediction interval was also calculated: the interval giving the distribution of effects across studies/liming operations/rivers, within which 95% of true effects were predicted to occur. Additionally, the percentage of true effects predicted to be negative was calculated for fish abundance, assuming a normal distribution of true effects of the log ratio.

Where there were sufficient data, the impacts of potential effect modifiers were tested including type of study, type of liming, presence of stocking and the mean length of time over which calcium carbonate was applied. Fish abundance provided the largest dataset and hence the greatest opportunity for the analysis of potential effect modifiers. The impact of liming was calculated as the average over all samples and years. Hence, the length of the liming operation in a study was calculated as the average number of years liming had occurred over all of the samples taken in the study (i.e. if a study presented data for each of the first 10 years of continuous liming the average length of liming over those 10 samples would be five years). The impact of liming was averaged across years to account for inter-annual variability that can naturally occur in populations. Catchment liming operations are often one-off applications that are expected to have effects over a prolonged period. Only one study (out of 33) with data on fish abundance used catchment liming, so this was excluded from the length of liming analysis.

Differences in the impact of liming on abundance could reflect variation in species composition among limed sites. Several studies presented the change in fish abundance from particular species of fish rather than the overall change in the abundance of all fish. The most common fish present were Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*), and effects on these two were analysed separately.

3. Results and discussion

3.1. Impact of liming on fish

Overall, liming of freshwater was accompanied by increased fish abundance (mean effect (log ratio) = 0.53, SE = 0.14, z = 3.63, p = 0.003, Fig. 1). However, the effect varied significantly among individual studies (random effects meta-analysis on log ratios, Qm = 363, df = 33, $p < 0.001$), and the prediction interval, in which 95% of true effects were predicted to occur, overlapped zero (log response ratio of −0.4 to 1.5). In other words, liming was predicted to increase fish abundance overall, but in any one river there was an 18% possibility of fish abundance decreasing after liming. An idea of what the ratio corresponded to in absolute terms was given by the baseline abundance estimates for the three most common measures: 36 (SD 25) fry and par per 100 m^2 (n = 17), 5.1 (SD 6.1) adult fish per 100 m^2 (n = 4) and 186 (SD 91) total kg caught per year (n = 5).

3.2. Factors affecting the impact of liming on fish

The experimental designs involved in liming as well as the water chemistry of rivers after liming varied among studies (Table 2). However, there was no significant difference in the effect of liming between studies of different designs, and study type was not a significant effect modifier within the random effects meta-analysis (Qm = 0.796, df = 1, p = 0.372). Hence, the study design did not appear to have caused the observed effects. Nor was the presence of stocking a significant effect modifier (Qm = 0.587, df = 1, p = 0.444). Liming method (river, catchment or lake) was also not a significant effect modifier (Qm = 0.559, df = 1, p = 0.455) in contrast to the average length of time the river had been dosed (Qm = 5.379, df = 1, p = 0.020, Fig. 2). Number of treatment years generally increased effect size (Fig. 2). However, the study with the largest effect size had involved liming for an average of two years (Shaver Fork study, Clayton et al., 1998). None of the rivers which had been limed for an average of more than 7.5 years showed a negative effect size.

The impact of the length of operation on the effectiveness of liming might reflect either the effects of prolonged chemical stability (e.g. Ormerod and Durance, 2009), or intrinsic population processes. Thus, populations with inter-annual life cycles will recover after disturbance over periods of years. For example, Atlantic salmon (*Salmo salar*) require on average 3–6 years to reach maturity depending on region (Hutchings and Jones, 1998). Recolonisation processes are also time-dependent. However, the

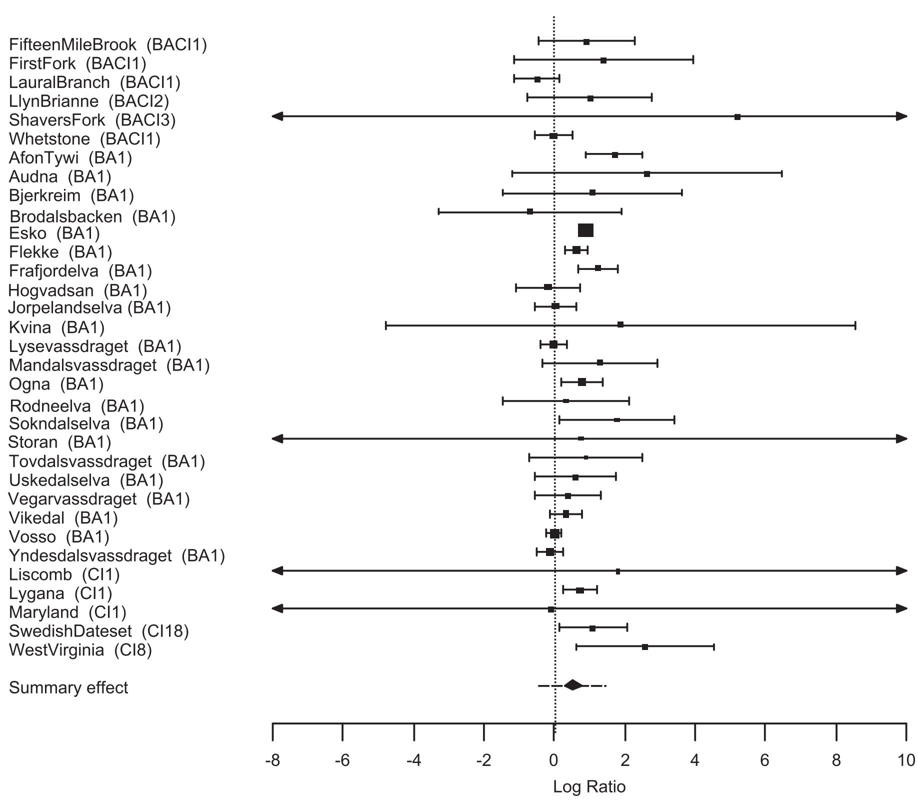

Fig. 1. Forest plot of fish abundance response to freshwater liming. Squares are the effect size for each study, error bars are the confidence interval for each study (arrows indicate a larger confidence interval (shavers CI −56 to 66) or where insufficient data were presented to calculate a confidence interval (Liscomb, Maryland and Storan)). Summary effect: diamond represents the weighted mean calculated from the random effects meta-analysis, width of the diamond is proportional to the estimation in the error of the mean and the dashed line is the prediction interval where 95% of true effects are predicted to lie. The study design and number of rivers included are in brackets after each study name.

Table 2
Details of the studies included in the fish analysis including study design, presence of stocking (Sto.), Water quality parameters and liming method. Details of the water quality include the mean pH from the control data and the mean pH, Al (mg/l) and Ca (mg/l) concentration in the treatment sites and years. na = the information was not available for the study. The liming details given are M – the liming method (D = doser, L = lake, C = catchment, P = point), T – the timing of liming (C = continuous, A = annual, 1 = once), L – the average length of liming in years and D – the dose of lime applied (V = varied, U = unclear from study, A = automatic depended on flow or pH).

Study name, location	Study design	Sto.	Water quality				Liming.			
			Control mean pH	Treatment			M	T	L	Dose
				pH	Al	Ca				
Afon Tywi, UK	BA1	U	5.6	6.6	0.22	2.3	DL	C	5	U
Audna, Norway	BA1	Y	5	6.3	0.05	1.8	DL	C	10.5	U
Bjerkreim, Norway	BA1	Y	na	na	0.01	1.0	DL	C	6	U total 2200–2700 t
Brodalsbäcken	BA1	U	na	na	na	na	DL	C	1	U
Esko, Norway	BA1	Y	na	6.6	0.06	1.8	D	C	6	V 421–792 t per yr
Fifteen Mile Brook, Canada	BACI1	N	4.7	5.0	0.16	1.3	P	P	3.5	320 t dolomitic limestone
First Fork, USA	BACI1	N	5.3	6.0	0.17	3.9	P	A	1.5	505 t limestone total
Flekke, Norway	BA1	Y	5.4	6.2	0.10	1.0	D	C	5.5	U
Frafjordelva, Norway	BA1	Y	na	6.4	0.08	1.3	DL	C	7	V U
Högvadsån, Sweden	BA1	U	na	na	na	na	CL	A	2.5	U
Jørpelandselva, Norway	BA1	Y	'Acidic'	5.9	0.01	0.7	L	A	6.5	V U
Kvina, Norway	BA1	N	4.9	6.3	0.01	1.5	DL	C	7.5	V U
Laural Branch, USA	BACI1	N	na	na	<0.03	2	D	C	1	~8.2 t limestone
Liscomb River, Canada:	CI1	U	4.9	na	na	na	P	1	1	180 t
Llyn Brianne, UK	BACI2	U	5.1	6.2	0.09	3.5	C	1	1	25–9 t per ha $CaCO_3$
Lygana, Norway	CI1	N	na	6.4	0.01	1.7	DL	C	11	V: 2004–08 1763–3073 t per yr
Lysevassdraget, Norway	BA1	N	na	6.4	0.01	1.1	DL	C	4.5	V: 2004–08 73–253 t per yr
Mandalsvassdraget, Norway	BA1	Y	na	6.3	0.01	1.2	DL	C	5.5	~5000 t per year
Maryland, USA	CI1	U	6.5	6.7	na	na	D	C	2	U.
Ogna, Norway	BA1	N	5.7	6.5	0.01	1.9	DL	C	9	A V:151–389 t per yr
Rødneelva, Norwegian	BA1	Y	5.4	6.5	0.06	1.4	DL	C	6	V. U
Shavers Fork, USA	BACI3	N	4.8	5.7	0.23	4.6	P	A	1	Total 733 t limestone
Sokndalselva, Norway	BA1	N	5	6.2	0.08	1.2	L	A	10	V: 673 -970t per yr
Storån, Norway	BA1	U	na	na	na	na	L	1	1	U
Swedish Database	CI18	U	<6	6.5	0.19	5.0	DLC	C	16.7	V 673–12803 t 50% CaO per river
Tovdalsvassdraget, Norway	BA1	Y	5.1;	6.3	0.01	1.5	DL	C	6	V: 4466–6407 t per yr
Uskedalselva, Norway	BA1	U	6.0	6.2	0.01	1.1	D	C	3.5	V: 42–85 t per yr
Vegårvassdraget, Norway	BA1	Y	5.0	6.5	0.01	1.9	DL	C	6.5	V 2004–08, 261–463 t per yr
Vikedal Norway	BA1	N	6.0	6.6	<0.01	1.6	D	C	11	U, A,
Vosso, Norway	BA1	Y	6.2	6.4	0.04	1.1	DL	C	5	U A, by water flow
West Virginia. USA	CI8	U	4.9	6.8	0.14	3.2	P	U	11	U
Whetstone Book, USA	BACI1	N	5.9	6.6	0.13	2.7	D	C	1.5	Total 56 t limestone
Yndesdalsvassdraget, Norway	BA1	U	5	6	0.01	1.9	D	C	7	U

confounding effect of amelioration in acid deposition over time cannot be ignored (Ormerod and Durance, 2009). As many of the studies did not have data from control locations, there was no control for the change over time occurring even without the presence of liming. Acidic deposition has been decreasing over time (Evans et al., 2001), and effects may have been more visible in longer studies. It is also worth noting that the majority of studies have been published since 1995 (38 out of 45), so will have been confounded by decreased acid deposition. Additionally, the scale of the liming operation was generally linked to the duration of the intervention. Small scale experimental set-ups were often only operating for a few years whereas large-scale national operations, e.g. those in Norway and Sweden, were often in operation for many years. The mean duration of application did not explain all of the variation, so other factors must have been important.

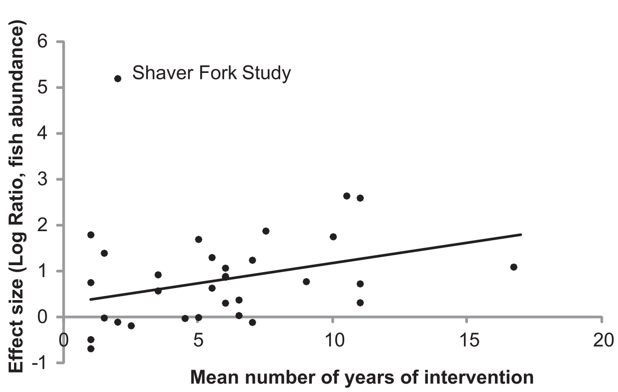

Fig. 2. The relationship between duration of liming intervention and effect of liming on fish abundance. The line represents a linear regression analysis omitting the Shaver Fork Study which was deemed an outlier.

3.3. Differences in liming impact on different fish species

Salmon populations increased to a greater extent than brown trout in all except one of the 19 rivers where effect sizes for both could be calculated (Fig. 3). The exception was the river Vosso (Direktoratet for naturforvaltning, 2009), where salmon density decreased after liming. Overall there was a significant mean increase in salmon (mean effect (log ratio) = 1.16, SE = 0.38, $z = 3.02$, $p = 0.003$) although there was significant heterogeneity between studies (Qm = 114, df = 18, $p < 0.0001$). For brown trout, where salmon were also present, there was no significant effect of liming. There was significant heterogeneity between studies (Qm = 302.3, df = 18, $p < 0.0001$), and the mean effect was small, non-significant and negative (mean effect (log ratio) = −0.15, SE = 0.17). The brown trout population, when salmon were also present, increased in some rivers after liming but decreased in others. However in each of the three studies where brown trout were studied, without the presence of salmon, the trout increased in abundance (Swedish database: log ratio = 0.79 (SE 0.5); Llyn Brianne, Weatherley and Ormerod, 1992: log ratio = 1.0 (SE 0.9); and Whetstone Brook, Simmons et al., 1996: log ratio = 0.6 (SE 0.2)).

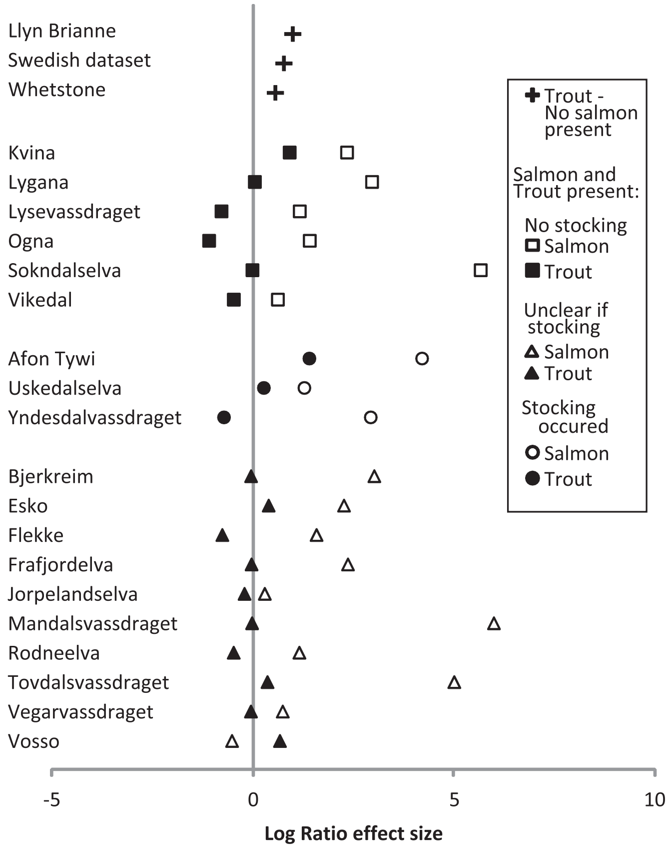

Fig. 3. Effect sizes for salmon and trout abundance in freshwater liming studies where both were measured. The black symbols are the effect sizes for trout and the white symbols represent the salmon effect sizes. The graph is also divided on the presence of stocking; triangles represent where there was no stocking, diamonds where stocking was not mentioned in the articles and squares where stocking took place.

Species do not live in isolation and interactions between species may result in indirect impacts on species abundance with liming. It has been suggested that trout numbers do not increase, or may even decrease, where salmon numbers increase, due to increased competition (Degerman and Appelberg, 1992). Hence, the fish community composition at the outset of liming may also be important for determining the impact, along with the order of any recolonisation (Clair and Hindar, 2005) and the presence of any stocking.

3.4. Limitations to the findings

Despite the apparent consistency of outcomes among studies with different designs, the fact that a minority of the fish abundance studies had both control and baseline data in a BACI design limits the confidence in the results. In studies with no control sites, factors other than liming may have changed over time and caused the observed differences. In the studies with no baseline data it was not possible to be certain whether the control and treatment groups were the same before liming occurred, and this is a basic and well-known problem in experimental design. Randomizing the assignment of each river to the control or treatment group decreases the chances of systematic differences between the control and treatment river, however, in no study was this apparent. The controls were either upstream sections of the limed river or in unlimed control rivers. Rivers and streams naturally vary along their length and hence differences in fish and invertebrate populations can be expected along the length of a stream. In the main Norwegian survey (Direktoratet for naturforvaltning, 2009), where control sites were present, they were often in unlimed tributaries which may have had systematically different characteristics to the main river, where the limed sites were located, not due to the liming. In smaller scale experimental studies the controls sites were generally closer (i.e. Little Stoney Creek; Downey et al., 1994 and Liscome; Watt et al., 1983) and hence may have been more similar. In studies where the controls were independent, unlimed rivers, in half of the studies (four out of eight) there was mention of some form of ‘matching’ or proximity of the control streams but it was not clear how this was done or how closely matched the sites were. However, in the other four studies the method of allocation was not mentioned. Also, in several of the studies the liming operations and sites were selected by those responsible for the operations and not the researchers carrying out the studies. Hence, differences observed in the control impact studies may be due to underlying differences in the rivers chosen to be limed.

Many of the studies within the Norwegian survey reported the use of stocking to aid the recovery of fish stocks from acidification. Such effects will alter fish abundance and, if stocking starts after liming, will be an important confounding factor. Nevertheless, there was not a significant difference in the results observed in studies where stocking occurred. Most studies outside Norway did not mention stocking. Inevitably given the global distribution of liming, a large proportion of the studies available occurred in Scandinavia and north western Europe so the results of the analysis are most relevant to liming activities in these regions.

3.5. Impacts of liming on acid sensitive invertebrates

In all studies that measured the abundance of acid-sensitive invertebrates (Table 3), liming was accompanied by an increase (Mean = 0.68, SE = 0.29, $z = 2.56$, $p = 0.018$). However, despite the effect not varying significantly between studies (Qm = 10.4, df = 5, $p = 0.06$, Fig. 4), the heterogeneity was still high (I^2, the percentage of total variability due to heterogeneity = 52%) and the prediction interval overlapped zero. There were insufficient data to exclude the possibility that abundance decreased in some rivers.

The number of acid sensitive taxa also increased with liming (mean log ratio = 0.95, SE = 0.23, $z = 4.16$, $p < 0.0001$), and this effect did not vary among studies (Qm = 0.52, df = 4, $p = 0.97$, Fig. 5). None of the variability was due to heterogeneity of studies ($I^2 = 0\%$). Two studies where log ratio effect sizes could not be calculated were not included in this analysis. In both (West Virginia and Dogway), the number of acid-sensitive invertebrate taxa was higher in the limed sites. In the West Virginia study (McClurg et al., 2007) there were no acid-sensitive invertebrates in the control and in the Dogway study (Menendez et al., 1996), the control increased from zero acid-sensitive invertebrates.

However, confidence in the mean calculated in the meta-analysis was driven largely by one study, Audna (Direktoratet for naturforvaltning, 2009), which had the smallest variance. After removal of this from the analysis, the mean effect was no longer significantly different from zero. Additionally, the Audna study, despite having the smallest variance, only covered one river, whereas the largest study, covering five rivers, (Herrmann and Svensson, 1995) was given the smallest weighting as the confidence interval could not be calculated.

3.6. Impact of liming on all invertebrates

Liming was not accompanied by any increase in overall invertebrate abundance. Over all studies, the mean liming effect was no change in abundance (mean effect (log ratio) = 0.01, SE = 0.12, $z = 0.07$, $p = 0.944$, Fig. 6) and the prediction interval, the interval

Table 3
Details of the studies included in the invertebrate analysis including study design, parameters covered by the study (SR = acid sensitive invertebrate richness, SA = acid sensitive invertebrate abundance, AR = all invertebrate richness, AA = all invertebrate abundance), water quality parameters and liming method. Details of the water quality include the mean pH from the control data and the mean pH, Al (mg/l) and Ca (mg/l) concentration in the treatment sites and years. Na = the information was not available for the study. The liming details given are: M – the liming method (D = doser, L = lake, C = catchment, P = point), T – the timing of liming (C = continuous, A = annual, 1 = once), and the dose of lime applied (V = varied, U = unclear from study, A = automatic depended on flow or pH).

Study name, location	Study design	Covers	Water quality				Liming		
			Control mean pH	Treatment			M	T	Dose
				pH	Al	Ca			
Audna, Norway	CI1	SR, AA	5	6.3	0.05	1.8	DL	C	U
Bear Run, USA	CI1	AR, AA	4.7	5.5	0.11	na	P	A	21–23 t per yr
Bjerkreim, Norway	CI1	SA	na	na	0.01	1.0	DL	C	U total 2200–2700 t
Dogway, USA	BAC1I	AR, AA	4.6	6.3	0.32	3.3	D	C	10 g/m^3
Esk, UK	BACI1	AA	5.2	5.8	0.07	1.8	C	2	3200 t limestone per yr
Herrmann, Sweden	BACI5	SR, AA	4.7	5.6	na	na	DL	C	V U
Larsson, Sweden	CI5	AA	4.7	4.8	na	na	C	1	V U
Laural Branch, USA	BACI1	AR	na	na	<0.03	2	D	C	~8.2 t limestone
Lingdell, Sweden	BACI12	AA	4.9	6.0	na	na	DLC	C	V U
Little Stoney Creak, USA	CI1	AR, AA	5.7	6.8	0.06	3.5	P	A	Total 105 t limestone
Llyn Brianne, UK	BACI3	SA, SR, AR, AA	5.1	6.2	0.09	3.5	C	1	25–9 t per ha $CaCO_3$
Loch Fleet, UK	BA1	AR	4.5	6.6	na	3.4	C	1	445 t limestone
Lygana, Norway	CI1	AR	na	6.4	0.01	1.7	DL	C	V: 2004–08 1763–3073 t per yr
Lysevassdraget, Norway	BACI1	SA	na	6.4	0.01	1.1	DL	C	V: 2004–08 73–253 t per yr
Mandalsvassdraget, Norway	BACI1	SA	na	6.3	0.01	1.2	DL	C	~5000 t per year
Mountain Run, USA	BACI1	AA	5.2	6.0	0.01	2.3	P	1	36 t limestone total
Ogna, Norway	CI1	SR	5.7	6.5	0.01	1.9	DL	C	AV 1999–2008,151–389t per yr
Olofsson, Sweden	BA2	AR			0.07	6.8	U	U	U
Parasites, Canada	CI1	AA	4.7	6.4	na	3.0	L	A	U
Pottsville, USA	CI5	AR	4.9	6.2	na	na	P	U	U
Sokndalselva, Norway	CI1	SA	5	6.2	0.08	1.2	L	A	V: 673–970 t per yr
Swedish Database	CI18	AR, AA	<6	6.5	0.19	5.0	DLC	C or A	V 673–12803 t 50% CaO per river
Vikedal Norway	CI1	SA, SR, AA	6.0	6.6	<0.01	1.6	D	C	U, A,
Vosges, France	BACI2	AR	4.7	5.1	0.50	1.7	C	1	Average 2.5 t per ha
West Virginia. USA	CI4	AR, AA	4.9	6.8	0.14	3.2	P	U	U
Whetstone Book, USA	BACI1	AR, AA	5.9	6.6	0.13	2.7	D	C	Total 56 t limestone
Wye, UK	CI3	AA	4.6	5.0	na	na	C	1	Total 750 t limestone

within which 95% of true effects are predicted to occur, ranged from −0.90 to 0.92 with significant variability between studies ($Qm = 117$, $df = 12$, $p < 0.0001$, Fig. 6). Hence, in some rivers, invertebrate abundance decreased after liming and in others increased. The liming method was not a good predictor of this variation ($Qm = 0.05$, $df = 1$, $p = 0.82$). However, there was a non-significant trend for the type of study ($Qm = 3.77$, $df = 1$, $p = 0.052$). The mean effect size for studies of BACI design was negative whereas the mean for before and after studies with no control was positive. The change in abundance in the latter could be due to other factors changing over time. This potential difference between the studies suggests that the risk of bias within the BA studies may be affecting the results. If the BACI studies are taken alone then the mean effect was a significant negative effect (mean effect (log ratio) = −0.24, SE = 0.12, $z = -1.99$, $p = 0.047$), although there was still significant variability between the studies ($Qm = 27.88$, $df = 5$,

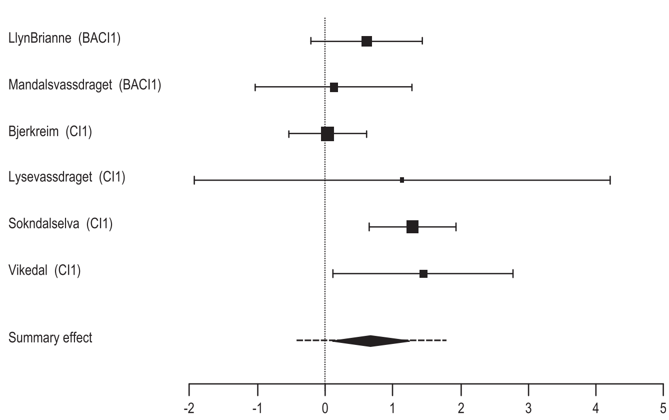

Fig. 4. Forest plot of acid sensitive invertebrate abundance effect size in response to freshwater liming for each of the studies. The squares are the effect size for each study, the error bars are the confidence interval for each study. Summary effect: the diamond represents the weighted mean calculated from the random effects meta-analysis, the width of the diamond is proportional to the estimation in the error of the mean and the dashed line is the prediction interval where 95% of true effects are predicted to fall. The study design and number of rivers included are in brackets after each study name.

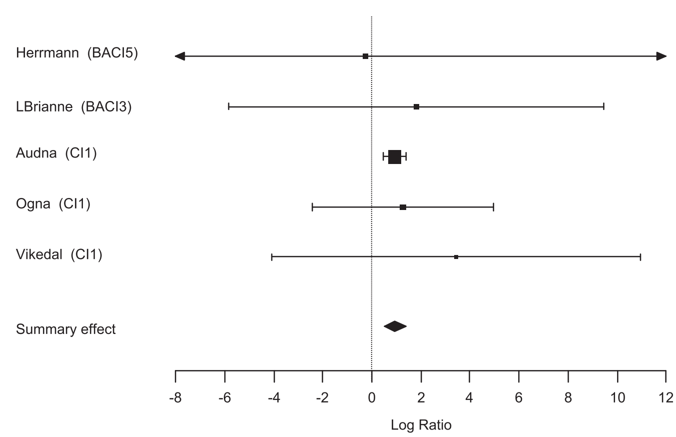

Fig. 5. Forest plot of the effect size for the number of acid sensitive invertebrate taxa in response to freshwater liming for each of the studies. The squares are the effect size for each study, the error bars are the confidence interval for each study (arrows indicate where insufficient data were presented in the study to calculate a confidence interval). Summary effect: the diamond represents the weighted mean calculated from the random effects meta-analysis and the width of the diamond is proportional to the estimation in the error of the mean. Unlike in the other plots, there is no dotted line extending from the diamond as the heterogeneity was calculated as zero. The study design and number of rivers included are in brackets after each study name.

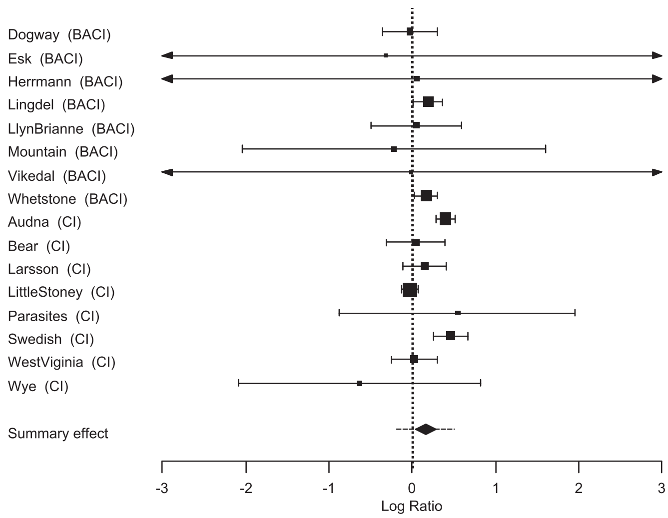

Fig. 6. Forest plot of invertebrate abundance in response to freshwater liming. The experimental design for each study is given in brackets after the study name. The squares are the effect size for each study, the error bars are the confidence interval for each study. Summary effect: the diamond represents the weighted mean calculated from the random effects meta-analysis, the width of the diamond is proportional to the estimation in the error of the mean and the dotted line is the prediction interval where 95% of true effects are predicted to fall.

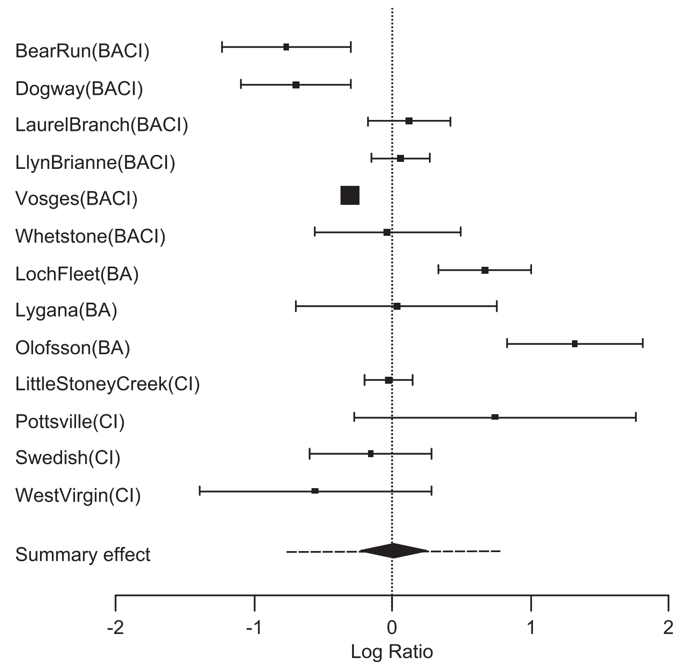

Fig. 7. Forest plot of invertebrate taxonomic richness in response to freshwater liming. The experimental design for each study is given in brackets after the study name. The squares are the effect size for each study, the error bars are the confidence interval for each study. Summary effect: the diamond represents the weighted mean calculated from the random effects meta-analysis, the width of the diamond is proportional to the estimation in the error of the mean and the dotted line is the prediction interval where 95% of true effects are predicted to fall.

$p < 0.0001$). Hence, invertebrate abundance may on average decrease after liming, although more studies of a BACI design are needed to be confident in this assessment.

The number of all invertebrate taxa present on average increased with liming (mean = 0.16, SE = 0.06, $z = 2.48$, $p = 0.013$), although the mean impact was smaller than for acid sensitive invertebrates alone (0.16 compared to 0.95). This effect varied significantly between studies (Qm = 42.8, df = 15, $p = 0.0002$, Fig. 7) and the prediction interval overlapped zero, so that some true effects were predicted to be negative. The type of study and liming method were not good predictors of this variation and the type of study (i.e. CI, BA or BACI) was not a good indicator of the study quality: although BACI studies generally used a high quality method, not all those used here were well designed. In three cases, the Esk and Vikedal from the Norwegian survey (Direktoratet for naturforvaltning, 2009) and Herrmann and Svensson (1995), no error could be calculated due to the lack of replication or a lack of reporting despite employing a BACI design.

3.7. Potential reasons for negative and varying impacts

In some instances liming itself may cause negative impacts on fish and invertebrates directly, for example due to changing the substrate or creating boundary conditions between limed and unlimed sites where changing pH might increase aluminium toxicity (Teien et al., 2006). Additionally, palaeolimnological studies have revealed that several limed lakes may have been naturally acidic, so that increased pH might have removed some naturally occurring taxa (Norberg et al., 2008). However, the decreases may not be related directly to liming. If other factors were limiting fish populations, for example other sources of pollution, fish populations could decrease despite any potential benefit from increased liming. The difficulty in isolating the impact of liming with studies lacking baseline and control data highlights the importance of future studies being well designed.

One of the key requirements in guiding where and when to use liming is understanding and predicting effectiveness. The duration of intervention and species present may be able to explain some of the variability but not all. The fact that the type of liming did not explain a significant amount of heterogeneity does not prove that the type of liming is not important or has no effect on the impact of liming. Instead, liming methods may not have explained a significant amount of the heterogeneity due to only having a small sample and other factors causing more heterogeneity and variation in effect. Factors that could not be investigated in this review but may be important include the dose of calcium carbonate applied, the community of organisms already present, the chemical conditions in the stream including the presence of acid episodes, changes in land use or management and changes in aluminium speciation when pH was increased by liming. The variation in effects, and the fact that other factors than liming cause a large variation in abundances, also suggest that if there is limited funding available all sources of population restrictions need to be considered. As acid deposition is decreasing, other factors impacting on the river may be of greater concern (e.g. barriers to migration, Hesthagen et al., 2011).

3.8. Wider impacts of liming

In the past, concerns about the effects of liming have been raised with respect to impacts on the conservation value of ecosystems that may be naturally base-poor. Liming effects will not be restricted to fish and invertebrates, and changes in the species richness and abundance of these organisms might reflect effects at lower trophic levels. Within Europe, the Water Framework Directive requires the identification of what constitutes ‘good ecological status’ for each type of water body. Different countries have traditionally used different indices to assess the ecological impact of acidification, including the AWIC (Acid Water Indicator Community) in the U.K., the Raddum index and the NIVA (Norwegian Institute for Water Research) index in Norway, (Moe et al., 2010).

Alongside these other indices, abundance and species richness are partial measures of the status of biological communities and ecosystems. Nine of the rivers in the Norwegian liming survey used the index developed by Raddum to measure the impact of liming on invertebrates, however as indices vary across Europe it was not used outside Norway. Hence, the indices could not be used as the measured outcome in this review. It is only relatively recently that more widely applicable tools have been developed (Moe et al., 2010).

The question of whether to lime is a complicated issue and only partly answered by the meta-analysis in this review. The review covered a limited question; "What is the impact of liming on fish and invertebrate richness and abundance?" Limiting the question in this way allowed the rigorous collection of all data on the topic and their statistical analysis. It showed that, on average, liming increased fish abundance and invertebrate taxon richness but in a minority of rivers fish abundance and invertebrates decreased. There are also multiple other aspects to the question of should we lime including the impact of liming on non-target habitats (including terrestrial habitats if catchment liming is implemented, Shore and Mackenzie, 1993), the cost of liming (Navrud, 2001) and the political and social reasons for liming (Clair and Hindar, 2005).

4. Conclusions

Liming was linked with increased fish abundance on average by 1.7 times the number of fish in the control sites or years. Salmon abundance increased in all except one river. Additionally, the mean abundance and taxonomic richness of acid sensitive invertebrates also increased along with the mean taxonomic richness of all invertebrates combined. Hence, on average liming appeared to be effective in mitigating two of the major ecological effects of acidification. However, these effects could not be guaranteed in all cases due to variability in outcomes among studies. Moreover, in several cases experimental design affected our ability to ascribe outcomes unequivocally to treatment effects. The mean increase in fish abundance and acid sensitive invertebrates with liming suggests that liming would on average be an effective mechanism for speeding up recovery. However, if decreased deposition engenders recovery independently, liming may not be necessary particularly where there are risks of negative impacts on fish and invertebrates.

Acknowledgements

Thanks to James Bussell and Diane Jones for their assistance in the early stages of this review. Also thanks to Barbara Livoreil for help with translating the French articles and to Tobias Vrede, Associate professor, Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences for compiling the Swedish dataset. Additionally, this work would not have been possible without the work conducted by the people who carried out the primary studies. We acknowledge funding from the UK Natural Environment Research Council through Knowledge Exchange Grant NE/F009356/1.

References

Barlaup, B.T., 2004. Vossolaksen – bestandsutvikling, trusselfaktorer og tiltak. DN utredning 2004–7. Direktoratet for Naturforvaltning, Norway.
Bostedt, G., Löfgren, S., Innala, S., Bishop, K., 2010. Acidification remediation alternatives: exploring the temporal dimension with cost benefit analysis. Ambio 39, 40–48.
CEE, 2010. Guidelines for Systematic Reviews in Environmental Management. Version 4.0. Collaboration for Environmental Evidence.
Clair, T.A., Hindar, A., 2005. Liming for the mitigation of acid rain effects in freshwaters: a review of recent results. Environmental Reviews 13, 91–128.
Clayton, J.L., Dannaway, E.S., Menendez, R., Rauch, H.W., Renton, J.J., Sherlock, S.M., Zurbuch, P.E., 1998. Application of limestone to restore fish communities in acidified streams. North American Journal of Fisheries Management 18, 347–360.
Degerman, E., Appelberg, M., 1992. The response of stream-dwelling fish to liming. Environmental Pollution 78, 149–155.
Donnelly, A., Jennings, E., Allott, N., 2003. A review of liming options for afforested catchments in Ireland. Biology and Environment: Proceedings of the Royal Irish Academy, Section B 103B, 91–105.
Downey, D.M., French, C.R., Odom, M., 1994. Low cost limestone treatment of acid sensitive trout streams in the Appalachian Mountains of Virginia. Water, Air, & Soil Pollution 77, 49–77.
Direktoratet for naturforvaltning, 2009. Kalking i laksevassdrag. Effektkontroll i 2008. Notat 2009-2.
Edwards, P., Clarke, M., DiGuiseppi, C., Pratap, S., Roberts, I., Wentz, R., 2002. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. Statistics in Medicine 21, 1635–1640.
European Union, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 Establishing a Framework for Community Action in the Field of Water Policy.
Evans, C.D., Cullen, J.M., Alewell, C., Kopácek, J., Marchetto, A., Moldan, F., Prechtel, A., Rogora, M., Veselý, J., Wright, R., 2001. Recovery from acidification in European surface waters. Hydrology and Earth System Sciences 5, 283–297.
Evans, C.D., Reynolds, B., Hinton, C., Hughes, S., Norris, D., Grant, S., Williams, B., 2008. Effects of decreasing acid deposition and climate change on acid extremes in an upland stream. Hydrology and Earth System Sciences 12, 337–351.
Firth, J.N.M., Ormerod, S.J., Prosser, H.J., 1995. The past, present and future of waste management in Wales: a case study of the environmental problems in a small European region. Journal of Environmental Management 44, 163–179.
Fowler, D., Smith, R., Muller, J., Cape, J.N., Sutton, M., Erisman, J.W., Fagerli, H., 2007. Long term trends in sulphur and nitrogen deposition in Europe and the cause of non-linearities. Water, Air & Soil Pollution: Focus 7, 41–47.
Herrmann, J., Svensson, B.S., 1995. Resilience of macroinvertebrate communities in acidified and limed streams. Water, Air, & Soil Pollution 85, 413–418.
Hesthagen, T., Larsen, B.M., Fiske, P., 2011. Liming restores Atlantic salmon (*Salmo salar*) populations in acidified Norwegian rivers. Canadian Journal of Fisheries and Aquatic Sciences 68, 224–231.
Hutchings, J.A., Jones, M.E.B., 1998. Life history variation and growth rate thresholds for maturity in Atlantic salmon, *Salmo salar*. Canadian Journal of Fisheries and Aquatic Science 55, 22–47.
Kowalik, R.A., Cooper, D.M., Evans, C.D., Ormerod, S.J., 2007. Acidic episodes retard the biological recovery of upland British streams from chronic acidification. Global Change Biology 13, 2439–2452.
Kuylenstierna, J.C.I., Rodhe, H., Cinderby, S., Hicks, K., 2001. Acidification in developing countries: ecosystem sensitivity and the critical load approach on a global scale. Ambio 30, 20–28.
Mant, R., Jones, D.L., Bussell, J., Godbold, D.L., Reynolds, B., Ormerod, S., Pullin, A.S., 2010. Is 'liming' of Streams and Rivers an Effective Intervention for Managing Water Quality to Support Fish and Invertebrate Populations? CEE Protocol 09–015. Collaboration for Environmental Evidence. www.environmentalevidence.org/SR76.html.
Mant, R., Jones, D.L., Reynolds, B., Ormerod, S., Pullin, A., 2011. What is the Impact of Liming of Streams and Rivers on the Abundance and Diversity of Fish and Invertebrates? CEE Review 09–015 (SR76). Collaboration for Environmental Evidence. www.environmentalevidence.org/SR76.html.
Matejko, M., Dore, A.J., Hall, J., Dore, C.J., Błaś, M., Kryza, M., Smith, R., 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.
McClurg, S.E., Petty, J.T., Mazik, P.M., Clayton, J.L., 2007. Stream ecosystem response to limestone treatment in acid impacted watersheds of the Allegheny plateau. Ecological Applications 17, 1087–1104.
Menendez, R., Clayton, J.L., Zurbuch, P.E., 1996. Chemical and fishery responses to mitigative liming of an acidic stream, Dogway Fork, West Virginia. Restoration Ecology 4, 220–233.
Menz, F.C., Seip, H.M., 2004. Acid rain in Europe and the United States: an update. Environmental Science & Policy 7, 253–265.
Moe, S.J., Schartau, A.K., Bækken, T., McFarland, B., 2010. Assessing macroinvertebrate metrics for classifying acidified rivers across northern Europe. Freshwater Biology 55, 1382–1404.
Moiseenko, T.I., 2005. Effects of acidification on aquatic ecosystems. Russian Journal of Ecology 36, 93–102.
Navrud, S., 2001. Economic valuation of inland recreational fisheries: empirical studies and their policy use in Norway. Fisheries Management and Ecology 8, 369–382.
Norberg, M., Bigler, C., Renberg, I., 2008. Monitoring compared with paleolimnology: implications for the definition of reference condition in limed lakes in Sweden. Environmental Monitoring and Assessment 146, 295–308.
Ormerod, S.J., Durance, I., 2009. Restoration and recovery from acidification in upland Welsh streams over 25 years. Journal of Applied Ecology 46, 164–174.
Reuss, J.O., Johnson, D.W., 1986. Acid Deposition and the Acidification of Soils and Water. Spirnger-Verlag, New York.
Reynolds, B., Stevens, P.A., Brittain, S.A., Norris, D.A., Hughes, S., Woods, C., 2004. Long-term changes in precipitation and stream water chemistry in small forest and moorland catchments at Beddgelert Forest, north Wales. Hydrology and Earth System Sciences 8, 436–448.

Sandøy, S., Langåker, R., 2001. Atlantic salmon and acidification in southern Norway: a disaster in the 20th century, but a hope for the future? Water, Air & Soil Pollution 130, 1343–1348.
Schindler, D.W., Mills, K.H., Malley, D.F., Findlay, D.L., Shearer, J.A., Davies, I.J., Turner, M.A., Linsey, G.A., Cruikshank, D.R., 1985. Long-term ecosystem stress: the effects of years of experimental acidification on a small lake. Science 288, 1395–1401.
Simmons, K.R., Cieslewicz, P.G., Zajicek, K., 1996. Limestone treatment of Whetstone Brook, Massachusetts. 2. Changes in the brown trout (*Salmo trutta*) and brook trout (*Salvelinus fontinalis*) fishery. Restoration Ecology 4, 273–283.
Shore, R.F., Mackenzie, S., 1993. The effects of catchment liming on shrews *Sorex* spp. Biological Conservation 64, 101–111.
Teien, H.-C., Kroglund, F., Salbu, B., Rosseland, B.O., 2006. Gill reactivity of aluminium-species following liming. Science of The Total Environment 358, 206–220.
Viechtbauer, W., 2010. Conducting meta-analyses in R with the metaphor package. Journal of Statistical Software 36, 1–48.
Watt, W.D., 1987. A summary of the impact of acid rain on Atlantic salmon (*Salmo salar*) in Canada. Water, Air, and Soil Pollution 35, 27–35.
Watt, W.D., Farmer, G.J., White, W.J., 1983. Studies on the use of limestone to restore Atlantic salmon habitat in acidified rivers. Lake and Reservoir Management 1, 374–379.
Weatherley, N.S., Ormerod, S.J., 1992. The biological response of acidic streams to catchment liming compared to the changes predicted from stream chemistry. Journal of Environmental Management 34, 105–115.