Metrics for evaluating the ecological benefits of decreased nitrogen deposition

Rowe EC1*, Jones L1, Dise NB2, Evans CD1, Mills G1, Hall J1, Stevens CJ1, Mitchell RJ4, Field C5, Caporn SJM5, Helliwell RC1, Britton AJ4, Sutton M3, Payne, RJ6, Vieno M1, Dore AJ1, & Emmett BA1

1 Centre for Ecology and Hydrology, ECW, Deiniol Road, Bangor, LL57 2UW, UK.
2 Centre for Ecology and Hydrology, Bush Estate, Penicuik, EH26 0QB, UK.
3 Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK.
4 The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK.
5 Manchester Metropolitan University, Manchester, M15 6BH, UK.
6 Environment, University of York, York, YO10 5DD, UK.
* Corresponding author. Email: ecro@ceh.ac.uk. Tel: +44 1248 374524.

Abstract

Atmospheric pollution by reactive nitrogen (N) can have profound effects on ecosystem functioning and biodiversity. Numerous mechanisms are involved, and response times vary among habitats and species. This complex picture can make it difficult to convey the benefits of controlling N pollution to policy developers and the public. In this study we evaluate pressure, midpoint, and endpoint metrics for N pollution, considering those currently in use and proposing some improved metrics. Pressure metrics that use the concept of a critical load (CL) are useful, and we propose a new integrated measure of cumulative exposure above the CL that allows for different response times in different habitats. Biodiversity endpoint metrics depend greatly on societal values and priorities and so are inevitably somewhat subjective. Species richness is readily understood, but biodiversity metrics based on habitat suitability for particular taxa may better reflect the priorities of nature conservation specialists. Midpoint metrics indicate progress towards desired endpoints – the most promising are those based on empirical evidence. Moss tissue N enrichment is responsive to lower N deposition rates, and we propose a new Moss Enrichment Index (MEI) based on species-specific ranges of tissue N content. At higher N deposition rates, mineral N leaching is an appropriate midpoint indicator. Biogeochemical models can also be used to derive midpoint metrics which illustrate the large variation in potential response times among ecosystem components. Metrics have an important role in encouraging progress towards reducing pollution, and need to be chosen accordingly.

Keywords: ammonium, global change, nitrate, nutrient, recovery.

Highlights:
Metrics are important for communicating progress in decreasing nitrogen (N) pollution
We evaluate pressure, midpoint, and endpoint metrics for N pollution
We propose new pressure metrics based on recent deposition above the critical load
Moss tissue N, and N leaching, are good midpoints at low, and high, N deposition
Biodiversity endpoints need to reflect societal values as well as natural science
Introduction

Atmospheric pollution by reactive nitrogen (N) is a global threat to biodiversity (Bobbink et al., 2010; Pardo et al., 2011; Phoenix et al., 2006; Sala et al., 2000) and is driving major changes in semi-natural habitats (e.g. Clark et al., 2013; Hauck et al., 2013; Song et al., 2012; Stevens et al., 2011a). Nitrogen availability often constrains plant growth (Elser et al., 2007), and although alleviating N limitation is of critical importance in agricultural systems (Ladha et al., 2005; Vanlauwe and Giller, 2006), the consequences of increased N deposition in more natural systems can be profound. Impacts can also be long-lasting because of N retention and recycling within the ecosystem, and because of depletion of seed banks (Basto et al., 2015) and delayed recolonisation. Efforts to decrease atmospheric N pollution need to be supported by an understanding among scientists and policymakers of the effects of present-day and historic emissions on ecosystems. Metrics have an important role in communicating the effects of policy decisions. We assessed current metrics used to represent benefits of decreases in N deposition, and propose new metrics to better represent nitrogen pressure and responses.

Many types of observations have been proposed as indicators of N pollution, such as plant tissue N concentration, litter C/N ratio, or plant species richness, but these are sometimes difficult to measure, not consistently related to the degree of pollution by N, or affected not only by N pollution but by management change and other drivers. A complicating factor is that N pollution is beneficial in some respects, not only as ‘free’ fertiliser for farmers and foresters but by increasing the fixation and storage of carbon (C) in woodlands, at rates estimated at 15-40 kg C kg⁻¹ N (de Vries and Posch, 2011). However, untargeted applications of N are inefficient and have unintended consequences. Overall assessments also need to take into account the major impacts of atmospheric N pollution on human health and on tropospheric ozone formation, but here we focus on metrics suitable for assessing the direct impacts of N on ecosystems. Metrics can:

- represent the pressure, defined as “physical expression of human activities that could change the status of the environment in space and time” (EEA, 2015), on the ecosystem;
- illustrate achievement of a desired endpoint, i.e. an aspect of the environment that is directly important and relevant to people. Examples are metrics that can be directly related to favourable conservation status, or that indicate attainment or failure of a water quality target;
- be seen as midpoints or “links in the cause-effect chain” (Bare et al., 2000) that represent progress towards or away from a desired endpoint, e.g. chemical conditions that make it likely that this endpoint will be achieved in future, or reductions in the abundance of a species that point to eventual local extinction.

The terms do not necessarily relate to the timescale of change, and ‘midpoint’ does not mean progress half-way towards a goal. The same metric may have a different role in relation to different targets – for example, the concentration of nitrate (NO₃⁻) in soil leachate is an endpoint metric for water quality since it is “of direct relevance to society’s understanding of the final effect” (Bare et al., 2000), but a midpoint indicator for biodiversity since it indicates progress towards changes in biological diversity.

Nitrogen affects terrestrial vegetation through direct toxic effects (especially on lichens and bryophytes), by increasing the growth of tall, fast-growing plants at the expense of shorter-growing and stress-tolerant species, and by the acidifying effect of nitrate leaching (Jones et al., 2014). Most evidence for biodiversity impacts is from studies on plants, although other taxa are affected via impacts on plants (Feest et al., 2014), in particular animals that require open microsites that may be shaded by increased vascular plant growth (Wallis de Vries and Van Swaay, 2006). Changes in plant tissue stoichiometry may also affect invertebrate herbivores directly (Vogels et al., 2013). Sensitive species can decline at very low absolute N deposition rates (Payne et al., 2013; Stevens et al., 2011c), or very low absolute ammonia (NH₃) concentrations (Cape et al., 2009). The form of N pollution can alter impacts on habitats, although whether it is oxidised or reduced N that is more damaging seems to be habitat-specific (van den Berg et al., 2016). Experiments on the effect of N form may have been influenced by effects on soil pH of the added counterion, and in any case the ratio of reduced to oxidised N in the soil environment is mainly determined by soil conditions and may differ greatly.
from the ratio in deposited N (Stevens et al., 2011b). Given these considerations, it seems adequate to consider total N flux as an indicator of N pollution pressure rather than NO\textsubscript{x} and NH\textsubscript{y} fluxes separately (RoTAP, 2012). By contrast, gaseous ammonia is phyto-toxic at much lower concentrations than nitrogen oxides and so needs to be considered separately. Nitrogen oxides also have an important role in the formation of ground-level ozone, harmful effects of which are reviewed elsewhere (e.g. Mills et al., 2016).

Air pollution policy makes extensive use of the concept of ‘critical load’ (CL), defined as “a quantitative estimate of an exposure to one or more pollutants below which significant harmful effects on specified sensitive elements of the environment do not occur according to present knowledge” (Nilsson and Grennfelt, 1988). Critical Load values for N have been defined on the basis of contribution to the acidity balance or of acceptable loss and immobilisation fluxes (Spranger et al., 2004). Another approach is to determine the CL using experimental and survey evidence regarding the N deposition rates at which biogeochemical or ecological changes begin to occur in different habitats, resulting in ‘empirical’ values (CL\textsubscript{empN}) (Bobbink and Hettelingh, 2011). The CL framework has been highly effective in driving reductions in sulphur pollution (Amann et al., 2011; Hordijk, 1991) and remains widely used in policy development.

Effects of N on ecosystems may be delayed by chemical buffering, and by delays in biological responses to the changed environment (Figure 1). As N deposition rate increases, declines in pH may be buffered by cation exchange or mineral weathering; and available N concentrations in soil solution may be buffered by increased immobilisation or by plant uptake. Plant nutrient uptake is a critical process in ecosystems, and biological responses may occur before discernable change in soil solution N concentration. Nevertheless, there are likely to be delays in biological responses to such chemical effects as changes in tissue stoichiometry. Organisms may persist for a time even in unfavourable environments. Conversely, organisms are often unable to immediately colonise a site where the environment has become more favourable, particularly where the species has become extinct in the locality. Limited or no recovery from N pollution has been observed in several studies where experimental treatments ceased (Power et al., 2006; Silvertown et al., 2006; Strengbom et al., 2001), although recovery has been observed in some cases (Královec et al., 2009). Reasons for variation in recovery responses are discussed further in Stevens (2016).
Figure 1. (Adapted from Posch et al., 2004). Delayed effects of changes in N deposition on a chemical indicator and a biological indicator in: a) a strongly-buffered ecosystem, and b) an ecosystem with limited buffering capacity. Deposition above the critical load causes a chemical response, for example in conditions in the soil solution, to exceed a critical level after time \((t_2 - t_1)\). The biological response to these conditions is further delayed, and only becomes critical after time \((t_3 - t_1)\), called the Damage Delay Time (DDT). Biological recovery after deposition declines below the critical load will similarly be delayed, by the Recovery Delay Time (RDT).

A good metric simplifies but still represents current scientific understanding, can be related to effects that are important to people, and is measurable or easily related to simple observations. In this study we discuss the relevance of proposed pressure, midpoint and endpoint indicators for summarising the dynamic impacts of N pollution on ecosystems.

**Pressure metrics**

In assessments of N pollution, the principal pressure metrics are those related to total N deposition rate, and to the atmospheric concentration of ammonia. As noted above, evidence that input fluxes of oxidised and reduced N need to be considered separately is limited, but gaseous ammonia represents a different type of pressure. Site-specific estimates of gaseous pollutant concentrations can be obtained using passive samplers (Puchalski et al., 2011; Sutton et al., 2001), but modelling approaches are usually more appropriate for site-scale flux estimates (Theobald et al., 2009). Atmospheric N concentrations and input fluxes are simulated using models of chemical reactions, transport and deposition, parameterised using data on emissions sources. Large-scale deposition models are calibrated and tested against observations of N concentrations in aerosols, precipitation and the gas phase from networks of monitoring sites (Dore et al., 2015), and have increasingly been resolved to finer spatial resolutions (Vieno et al., 2014). In this section we assess metrics for quantifying N pollution pressure, including deposition rates in relation to the CL.

A widely-reported metric of ecosystem damage, the percentage of sensitive habitat area where the CL for nutrient N is exceeded \((SA_{ex})\), is rather insensitive to decreases in N deposition, principally because CL is substantially exceeded over large areas. At European scale, \(SA_{ex}\) is likely to decrease only marginally by 2050 despite a forecast 67 % decrease in deposition (Simpson et al., 2014). The unresponsiveness of \(SA_{ex}\) is in part because this metric does not consider degrees of damage above the CL. Nitrogen impacts are progressive, and species may be lost with marginal increases in N deposition from rates that are already well above the CL (Emmett et al., 2011; Stevens et al., 2011c). Sensitive species can also decline at deposition rates below CL values as currently set (Armitage et
al., 2014; Henrys et al., 2011; Payne et al., 2013), although such evidence may argue for a reduction in CL in certain habitats, since the CL is designed to protect the most sensitive component of the ecosystem. An aggregated metric which incorporates the degree of exceedance is the average exceedance of CL$_{\text{cum}N}$ for habitats within a grid square, weighted by the habitats’ areas, termed Average Accumulated Exceedance (AAE), (Spranger et al., 2004).

Both AAE and SA$_{\text{cum}N}$ are based on current deposition, and do not take into account the persistence of pollutant N within ecosystems. Empirical evidence from systems that have received substantial additions of N without comparable increases in N loss fluxes (Moldan and Wright, 2011), together with modelling studies (Tipping et al., 2012), imply that pollutant N persists in soil and contributes to a long-lasting increase in the flux of mineralised N. This means that N impacts depend on historic as well as current deposition. Cumulative N deposition incorporates the duration as well as the rate of N input, and may be a better predictor of ecosystem impacts than is current deposition (Figure 2) (see also De Schrijver et al., 2011; Dupré et al., 2010; Phoenix et al., 2012).

Figure 2. Relationships between plant species richness in a survey of UK semi-natural habitats (recalculated from Field et al., 2014; Stevens et al., 2004) and: A) current N deposition, and B) cumulative nitrogen deposition since 1900. Deposition calculations are described in Payne (2014).

Calculating cumulative N deposition as total deposition over a long time period has several disadvantages. The historic spatial pattern of deposition is poorly known, and is often assumed to have been constant, resulting in a cumulative deposition map that has no more explanatory power than the current deposition map. Cumulative N deposition since a fixed date can only increase, but N deposited many decades previously is mainly unavailable to plants due to immobilisation into organic matter. Unless this immobilised N is released, due for example to a temperature-induced increase in mineralisation, it will have less biological impact than recently deposited N. Observed effects of changes in N deposition rate can be rapid (Bredemeier et al., 1998), particularly for sensitive bryophytes and lichens that interact primarily with atmospheric deposition onto foliar surfaces (Mitchell et al., 2004). A compromise between using cumulative total deposition and current deposition, which may respectively overemphasise and underemphasise the effects of persistent N, would be to calculate deposition above a threshold and for a relevant time period (Figure 3). A suitable integration threshold would be the amount of N that an ecosystem can process without harmful effects, which is the basis for the ‘steady-state mass balance’ approach to calculating CL (Hettelingh et al., 1995). Pre-industrial ecosystems would have received N from fixation and from the formation of oxidised N in lightning strikes, probably similar to the rate of 3-5 kg N ha$^{-1}$ yr$^{-1}$ estimated for unpolluted boreal systems by DeLuca et al. (2008). Some N is effectively lost from ecosystems through leaching, gaseous release, or long-term immobilisation into soil organic matter: net losses in unimpacted systems are estimated at 3-12 kg N ha$^{-1}$ yr$^{-1}$, the higher values mainly for woodland (Hall...
et al., 2003). The latter values are similar to CL_{empN} values, which have been defined for many habitats on the basis of empirical evidence (Bobbink and Hettelingh, 2011; Pardo et al., 2011). Although CL values are inevitably uncertain due to the difficulty of measuring N fixation and denitrification fluxes (in particular) and of characterising long-term effects, CL_{empN} values were set after extensive discussion among air pollution experts, and provide a good basis for an integration threshold.

Figure 3. Dependence of cumulative deposition on the exceedance threshold above which deposition is integrated, and on the integration period: e.g. 3 years preceding the present day (CE_{3}, vertical hatching), and 30 years preceding the present day (CE_{30}, horizontal hatching).

The most suitable start date for integrating deposition depends on the turnover rate of N in the ecosystem and thus the time for which deposited N remains active. Modelling and N recovery studies suggest that extra N will be retained in soil for extended periods (see below) and continue to become plant-available, albeit in gradually diminishing amounts. In epiphytic and epilithic ecosystems a relatively small substrate volume can be accessed by the flora (Crittenden, 1989), at least until substantial canopy necromass has accumulated (Nadkarni et al., 2004), and so N concentrations and substrate pH are likely to be buffered much less than in a soil-based system. We propose that N is likely to remain substantially active for an average of approximately 30 years in soil-based ecosystems (cf. Balesdent et al., 1988) and 3 years in epiphytic and epilithic ecosystems (cf. Clark et al., 2005; Jones, 2005; Mitchell et al., 2004), and that cumulative exceedance calculated over equivalent periods (CE_{3} and CE_{30}, respectively) are appropriate pressure metrics for these two types of ecosystem. These are illustrative values with a limited empirical basis, although they could be refined by isotopic tracing, and this is an important topic for further research. Decreases in deposition will decrease the CE_{30} and CE_{3} metrics immediately to an extent, and if maintained at a low level the cumulative deposition within the preceding timeframe will reduce commensurately.

Using different periods and thresholds for calculating cumulative deposition has implications for metric reporting. Where the same trajectory of ratios to current deposition is applied across a region, the spatial pattern of cumulative total deposition (e.g. Figure 4b) is identical to that of current deposition (Figure 4a). Integrating deposition above a threshold (Figure 4c) results in a larger proportion of the area being included in the lowest category than does integrating total deposition, and substantial areas of western and northern Britain are shown to have received comparatively little recent deposition above CL_{empN}. The hotspots of deposition shown in similar locations and with similar colours in Figures 4b and 4c, but these hotspots contrast with less-affected areas rather more clearly in Figure 4c.
**Figure 4.** Spatial patterns of total N deposition to UK dwarf shrub heathland calculated using the CBED model: a) recent deposition flux (annual mean 2004-6); b) cumulative total deposition 1970-2005; c) \( CE_{30} \), cumulative deposition over the critical load for nutrient N for dwarf shrub heath, 10 kg N ha\(^{-1}\) yr\(^{-1}\), in preceding 30 years (1986-2005). Temporal patterns of deposition were derived from Matejko et al. (2009). Data for all maps were subsetted using equal intervals on a linear scale.

**Midpoint metrics**

Midpoint metrics that represent progress towards or away from biodiversity endpoints are somewhat controversial, since it can be argued that any change in an ecosystem is directly relevant to biodiversity. According to the Habitats Directive of the EEC, a habitat is considered to have favourable conservation status when “the specific structure and functions which are necessary for its long-term maintenance exist” (EEC, 1992), and a change in any chemical variable within any organism or ecosystem pool could be seen as a change in function. However, chemical changes that require analytical equipment to discern are not immediately relevant to public perceptions of biodiversity, even if they provide mechanistic indications of the trajectory of the ecosystem. Conversely, changes in organisms that are sensitive to N but not important components of biodiversity could be seen as midpoint rather than endpoint indicators, and lichens in particular have been proposed for low-cost monitoring of N pollution (van Herk, 1999; Wolseley et al., 2015). To avoid extensive debate about which aspects of the chemical environment, and which organisms, are “directly important and relevant to people” (see Introduction) we will restrict discussion of midpoint metrics to chemical indicators, and discuss organismal changes in the following section on endpoint indicators. In this section we assess the utility of N stocks, concentrations and stoichiometry in plant tissue and soil; conceptual and modelled pools of N; and N loss fluxes, as midpoint indicators.

Nitrogen concentration in plant tissue has been shown to increase with N deposition in several gradient studies (e.g. Dise et al., 1998; Harmens et al., 2011) as well as in many experiments (e.g.
Jones, 2005; Lamers et al., 2000), although a survey by Aber et al. (2003) found no relationships between N deposition and foliar N in a deciduous and a coniferous tree species. In a review of ten long-term N-addition field experiments across several habitats, Phoenix et al. (2012) found tissue N concentration increased in either higher or lower plants, or both, in every experiment. Plants translocate N from leaves before senescence (Chapin III et al., 2012) so N limitation and demand within the ecosystem may be better reflected by N concentration in leaf litter than in live tissue. Litterfall N concentration was found to be the best predictor of N deposition rate, among those tested, in a survey of European forests (Dise and Gundersen, 2004).

Some lichen and bryophyte species are very physiologically sensitive to atmospheric N, particularly high gaseous or aerosol N concentrations (Cape et al., 2009; Sparrius, 2007), and bryophyte N concentration often increases with N deposition even at lower ranges of deposition (Mitchell et al., 2004; Pitcairn et al., 2006). Different species may have a different characteristic N content at any given N deposition level, and the saturation level is also species-specific (Figure 5a). A set of bryophytes is monitored in the European Moss Survey (Harmens et al., 2011; Harmens et al., 2014), and response functions for the response of moss tissue N to N deposition have been fitted. However, bryophytes can vary considerably in their responses to N deposition (Schroder et al., 2010; Stevens et al., 2011c). Information may be lost when deriving a response curve from data for several species, but species-specific responses would only be useful within the range of the species. For this reason we propose a simple metric, termed the 'Moss Enrichment Index' (MEI), in which tissue N concentration is normalised to a value between 0 and 1 (Equation 1).

\[
MEI = \frac{\%N_{\text{observed}} - \%N_{\text{minimum}}}{\%N_{\text{maximum}} - \%N_{\text{minimum}}}
\]  

(Equation 1)

where \(\%N_{\text{minimum}}\) and \(\%N_{\text{maximum}}\) represent the lowest and highest levels of tissue %N recorded for the species across a sufficiently broad gradient of N deposition (Figure 5b). The MEI has the advantage of providing a directly measurable, single metric of N enrichment within the ecosystem, which can be expected to respond relatively rapidly to changes in N deposition, and which may provide an indication of recent ecosystem N exposure at lower N deposition levels, for which other biogeochemical measurements such as mineral N leaching may be ineffective.
Figure 5. a) Moss tissue N plotted against current N deposition (kg N ha\(^{-1}\) yr\(^{-1}\)) for four mosses: *Racomitrium lanuginosum* (Raco) *Dicranum scoparium* (Dicr), *Isothecium myosuroides* (Isot) *Thuidium tamarascinum* (Thui). Data from: Jones (2005); Baddeley et al. (1994); Jonsdottir et al. (1995); Pearce & van der Wal (2002); Pearce et al. (2003); Leith et al. (2008); Armitage et al. (2012). b) The same data, normalised to a range from the minimum to maximum measured tissue N concentration for each species, to derive a Moss Enrichment Index, MEI. The curve shown, $\text{MEI} = 1 - e^{(-0.0323 \times \text{N deposition})}$, was fitted by minimising total sum of squared differences.

Since the C concentration in dry plant tissue is relatively uniform, plant tissue C/N ratio is approximately equivalent to N concentration and will not be considered separately here. Stoichiometries with respect to other elements may however be useful. Tissue N/P ratios are thought to reflect relative P limitation (Koerselman and Meuleman, 1996), and were observed to increase with N additions at three heathland sites in the review by Phoenix et al. (2012). However, a gradient study of *Calluna vulgaris* tissue chemistry showed greater N concentration with more N deposition, but an even greater proportional increase in tissue P concentration presumably because N stimulated P uptake (Rowe et al., 2008). This suggests that plant tissue N/P ratio is not a robust indicator of ecosystem responses to N deposition and recovery.

Ecosystems can retain large amounts of deposited N, much of it in soil N pools with slow turnover rates (Nadelhoffer et al., 1999). Heathland soils have been observed to retain remarkably large amounts of N in litter and organic upper soil horizons, even after 10 years of N addition at rates up to 120 kg ha\(^{-1}\) yr\(^{-1}\) (Pilkington et al., 2005). Grassland and bog soils appear to be less effective as long-term stores of N (Phoenix et al., 2012), although changes in N stock are inherently more difficult to detect in such soils since they are often spatially heterogeneous and stocks are large in relation to pollutant N inputs. Changes in soil N concentration or total C/N ratio are in principle easier to detect, although the issue still remains that the signal may be diluted by a large existing stock or masked by spatial variation (Moldan et al., 2006). It is often assumed that N retention will decrease soil C/N ratio (e.g. Aber, 1992; Mulder and et al., 2015), but N deposition may also stimulate the production and incorporation of plant litter with relatively high C/N ratio, causing increases in soil C/N ratio in some habitats (Jones et al., 2004; Reynolds et al., 2013). Changes in C/N ratio were not observed in an N-gradient study of European conifer forests (Dise et al., 1998), nor in a survey of UK acid grasslands (Stevens et al., 2006). The direction of change in C/N ratio induced by increased N deposition will depend on the degree to which N limits plant growth in the system, with increases where litter production is stimulated and decreases where immobilisation into soil N is the more significant process, and so soil C/N ratio is not reliable as a midpoint indicator.
The stock or concentration of plant-available N in soil is in principle a better indicator of N status than total N. The KCl-extractable mineral N concentration has been shown to be related to N deposition rate in experiments on upland heath, some grasslands and to a lesser extent at a bog site (Phoenix et al., 2012), and also in regional surveys of acid grassland (Stevens et al., 2006) and upland heath (Southon et al., 2013). Mineral N concentrations in litter in an upland heath fell after a decrease in experimental N addition (Edmondson et al., 2013). Plants can also use small organic molecules as sources of N (Hill et al., 2011), but there is little evidence that dissolved organic N concentration in soil extractions or leachate is a reliable indicator of N status. Nitrogen in soil solution is likely to fluctuate rapidly in relation to rainfall and mineralisation events, and rapid plant uptake and/or immobilisation into soil organic matter can lead to zero measurements even where the flux into plants is evidently non-zero (Schimel and Bennett, 2004). The plant-available pool is thus not straightforward to define or measure. Time-integrated measures such as resin-sorbed N or mineralisable N provide a better indicator of N status than instantaneous measurements (Schimel and Bennett, 2004), and mineralisable N has been shown to increase in organic soils along a large-scale N-deposition gradient (Rowe et al., 2012). However, there is little consensus on measurement methods, which limits the evidence base for determining relationships between these measurements and N deposition. Due to these difficulties, neither instantaneous nor time-integrated measures of plant-available N can be recommended as midpoint metrics across ecosystems.

Soil N compounds have different timescales of availability. Soluble ions and molecules are in principle immediately available to plants, lichens and soil microorganisms, although species vary in the N forms they can process, and uptake also depends on organisms having access to these soluble N compounds before they are leached. Soluble N held electrostatically on clay and organic matter surfaces will be released if the solution is depleted by plant uptake or leaching, so can be seen as part of the plant-available pool. The majority of N in soil cannot readily be taken up by plants and other organisms since it is either incorporated in larger organic molecules or inaccessible within soil aggregates. Some of this N is readily released, but organic matter that is protected within soil aggregates or sorbed to clay particles can persist for many years (Schmidt et al., 2011). The continuum of availability timescales is typically represented in dynamic soil models using discrete pools with characteristic turnover rates (e.g. Coleman et al., 1997; Parton et al., 1988). Such models can be used to illustrate the varying timescales of impacts, with rapid responses of soluble N to changes in deposition, but also accumulation of N in more stable soil pools and re-release from these pools over an extended period. For example, Figure 6 shows the effects of a hypothetical abrupt episode of N deposition as simulated using the MADOC model (Rowe et al., 2014) for a wet heath site (Migneint, UK: 52.993 °N, 3.813 °W), which uses conceptual organic matter pools with mean residence times at 10 °C of 2 years (‘fast’), 20 years (‘slow’) and 1000 years (‘passive’). The pools in this figure were normalised to a maximum of one; in fact the ‘passive’ N pool is around 700 times larger than the amount of plant-available N in a given year and the ‘slow’ N pool is around 10 times larger. It is difficult to test such long-term predictions, but the underlying N14C model has been calibrated using 14C dating to track the development of soil organic matter pools in the 12000 years since deglaciation (Tipping et al., 2012).
Figure 6. Increases in soil N pools (extra over constant low deposition scenario) with different turnover rates to a hypothetical increase in N deposition from 2 kg ha\(^{-1}\) yr\(^{-1}\) to 20 kg ha\(^{-1}\) yr\(^{-1}\) for the period 1970-2000, as predicted by the MADOC model for a peatland system.

Nitrogen loss fluxes from ecosystems can mainly be viewed as midpoint metrics, although nitrate concentrations in drinking water are directly relevant to environmental standards and so are also an endpoint metric. Nitrogen loss occurs even in unpolluted ecosystems, in particular through leaching of dissolved organic N (DON) which may determine long-term rates of net N accumulation (Vitousek et al., 2010). However, increases in loss fluxes indicate that the ecosystem is becoming saturated (Aber et al., 1998; Emmett, 2007). Denitrification fluxes have not been shown to be consistently related to experimental N addition rates (Phoenix et al., 2006), but nitrate leaching increases with experimental N addition at moderate to high N loads of 20-140 kg N ha\(^{-1}\) yr\(^{-1}\) (Dise and Wright, 1995; Phoenix et al., 2012) and a decrease in N load can lead to a rapid reduction in NO\(_3^-\) leaching (Boxman et al., 1998). Spatial patterns of NO\(_3^-\) in surface waters can be explained by N deposition rates (Allott et al., 1995). The rate of N leaching is not easy to measure directly within soil, but monitoring of surface-water nitrate can provide a robust and low-cost measure of changes in N status at catchment scale (provided there is no fertiliser use within the catchment). For this reason, and because the relationship between nitrate leaching and N deposition rate is reasonably consistent at least for sites with deposition rates > 25 kg N ha\(^{-1}\) yr\(^{-1}\), nitrate leaching flux can be considered a good midpoint metric for N pollution and recovery. Ammonium (NH\(_4^+\)) leaching is rarely observed since ammonium ions are sorbed relatively strongly onto soil surfaces (Phoenix et al., 2006), and ammonium reaching surface waters is likely to be rapidly nitrified. Although higher DON concentrations have been observed in leachate from dune (Jones et al., 2002), forest (Vanguelova et al., 2010) and heathland (Edokpa et al., 2015) ecosystems impacted by N, there is as yet insufficient evidence to recommend leaching fluxes of other forms of N as midpoint metrics.

Nitrogen leaching can lead to acidification. The acidification potential of deposited N depends on transformations in the soil, in particular on the amount that ends up being leached as nitrate (Reuss
and Johnson, 1986). In experiments both increases and decreases in pH have been observed (Phoenix et al., 2012), often due to the basic cations or acidic anions used as counter-ions to the added NH$_4^+$ or NO$_3^-$ (Evans et al., 2008). The value of pH as a metric of N pollution and recovery is in any case diminished by the impacts of historical sulphur deposition, which caused widespread and persistent acidification (Evans et al., 2014). Due to a dramatic fall in sulphur deposition since the 1970s, soil pH has since increased in some areas (Oulehle et al., 2011; Reynolds et al., 2013), which in turn is thought to have affected the N cycle (Kopacek et al., 2013). This consideration means that pH is not recommended as a midpoint metric for assessing N pollution.

**Endpoint metrics**

Nitrogen pollution has considerable direct and indirect effects on human health, water quality, and greenhouse gas fluxes, but these are well-reviewed elsewhere (e.g. Sutton et al., 2011). Here we focus on biodiversity endpoint metrics. Biodiversity can be seen in terms of diversity of various taxon groups, ‘habitat integrity’, similarity to a target or reference habitat, avoided extinction, ecosystem service provision, or from a host of other perspectives. Species richness is simple to measure and calculate, and it has been shown to be negatively correlated with current N deposition rate in acid grassland, heathland, sand dune and bog ecosystems (Field et al., 2014; Maskell et al., 2010; Stevens et al., 2011a). Species richness can be useful for translating N deposition scenarios into a term that is widely understood, and easily related to many conservation targets.

Simply counting the number of species can however mask large and potentially unfavourable changes in habitats (Curran et al., 2011). Species richness can increase with N pollution (Pierick et al., 2011), due to invasion by more eutrophic species (Roth et al., 2013). Such species are generally not targets for conservation, whereas small-growing species of oligotrophic environments tend to have higher threat status or be already locally extinct (Hodgson et al., 2014). Considering species richness within particular functional groups would allow better understanding of the underlying trends.

Individual species often provide important ecosystem functions and services, such as maintaining pollinator populations or having strong visual appeal, as well as being directly relevant to some definitions of biodiversity and closely linked to conservation targets. Nitrogen sensitivity does not per se imply importance to biodiversity endpoints, although in practice the more N-sensitive species are often of more conservation concern (Hodgson et al., 2014). Scarce species are a focus for nature conservation, but are not often used for habitat assessment since they are usually absent, and for the same reason their habitat-suitability niches are difficult to characterise. Methods for identifying species that indicate favourable habitat condition have been developed (e.g. Arponen et al., 2005; Landi and Chiarucci, 2010) and lists of target species proposed (e.g. Delbaere et al., 2009). The occurrence of such species, or their modelled habitat-suitability (Henrys et al., 2015), could be used as an endpoint metric of N impacts. Species that are distinctive for the habitat but not necessarily scarce may be a more suitable basis for biodiversity metrics (Rowe et al., 2016), and a “Habitat Suitability Index” (HSI) based on modelled habitat suitability for such ‘species of interest’ was recently adopted as a common metric for responses to the Working Group on Effects of the Convention on Long-Range Transboundary Air Pollution (Posch et al., 2014).

Several potentially-useful metrics can be derived from the traits of the species present, such as growth-form (e.g. shrub vs. herb, or graminoid vs. forb), physiology (e.g. typical specific leaf area or typical height), ecological strategy (e.g. competitive or stress-tolerant) or environmental preference. In Europe, environmental preference has often been expressed using ‘Ellenberg’ scores assigned to each plant species (Ellenberg et al., 1992; Hill et al., 2000). In a study based on large-scale survey data, mean values for several traits were shown to be sensitive to N deposition in at least some habitats: grass/forb cover ratio; Ellenberg N score (an indicator of productivity; Hill and Carey, 1997); mean Ellenberg R score (an indicator of alkalinity); mean typical canopy height; and mean typical specific leaf area (Emmett et al., 2011). In grasslands, the ratio of cover of grasses and forbs (i.e. non-grass herbs) was shown to be very responsive to N deposition load (Stevens et al., 2009). This relationship could be used to develop a responsive metric for these habitats. Sutton et al. (2009) proposed an index
derived from scores assigned to lichen species on the basis of their preference (or not) for acid and N-rich conditions. This “acidophytes / nitrophiles index” could be applied as an endpoint metric.

Conclusions and recommendations

The metrics most suitable for evaluating benefits of decreases in N deposition are summarised in Table 1. These all refer to total N deposition, since although there are differences between oxidised and reduced N in terms of potential controls on pollutant emissions, atmospheric transport and other processes, there is limited evidence that input fluxes of NO$_x$ and NH$_3$ have differing effects on habitats. The area where CL is exceeded, SA$_{ex}$, is relatively unresponsive to decreases in N deposition, reflecting the severe and ongoing damage caused by N. Conversely, the spatial average of exceedance, AAE, is more responsive to decreases in N deposition and reflects progress towards reduced damage. However, both of these pressure metrics are instantaneous measures and take no account of chemical and biological recovery delays. Midpoint indicators are more able to capture at least chemical delays to recovery. The most promising are tissue N concentration in mosses (for low-deposition systems) and N leaching (for high-deposition systems). These indicators vary in their responsiveness at different stages of ecosystem saturation with N, and are complementary in that moss tissue N concentration increases with reasonable consistency in the range 0-25 kg N ha$^{-1}$ yr$^{-1}$, whereas N leaching increases when N deposition is above this range. It may be more difficult to reach consensus on appropriate endpoint indicators for biodiversity, but species-richness and the HSI are complementary in that the former is easily understood, but the latter gives a more nuanced indication of habitat quality.

Table 1. Recommended metrics, classified by Type: P = Pressure; M = midpoint; E = endpoint.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Type</th>
<th>Appropriate for</th>
<th>Recommended calculation method</th>
<th>Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAE: Average Accumulated Exceedance</td>
<td>P</td>
<td>All habitats. All deposition rates above Cl$_{sat}$N.</td>
<td>Exceedance of Cl$_{sat}$N averaged across N-sensitive habitats within a grid-square, weighted by habitat area.</td>
<td>Pros: responsive and simple; ready to use. Cons: takes no account of impact delays.</td>
</tr>
<tr>
<td>CE$<em>3$ or CE$</em>{30}$: Cumulative exceedance</td>
<td>P</td>
<td>All habitats. All deposition rates.</td>
<td>Integrated exceedance of habitat-specific Cl$_{sat}$N, over the preceding 30 years for soil-based habitats or 3 years for epiphytic/epilithic sub-habitats.</td>
<td>Pros: responsive; well-related to timescale of impacts and to agreed definitions of damaging deposition rate. Ready to use. Cons: timescales based on expert judgement.</td>
</tr>
<tr>
<td>Moss Enrichment Index (MEI)</td>
<td>M</td>
<td>Habitats with mosses. Deposition rates up to 25 kg N ha$^{-1}$ yr$^{-1}$.</td>
<td>Measure moss tissue N % and compare with the N % range observed in the moss species, e.g. using relationships from Harmens et al. (2011).</td>
<td>Pros: well-correlated with (lower) deposition rates, easily measurable, useful ‘early warning’ metric. Cons: establishing data for new species requires data from sites with a range of N deposition</td>
</tr>
<tr>
<td>Stored N</td>
<td>M</td>
<td>Habitats with soil$^1$. All deposition rates.</td>
<td>Calculate ‘slow’ N pool in response to time-series of deposition using e.g. the N14C model (Tipping et al., 2012).</td>
<td>Pros: illustrates well a stock of N which places the habitat at risk; modelled values are easily upscaled. Cons: measurement methods remain uncertain.</td>
</tr>
<tr>
<td>N leaching rate</td>
<td>M/E</td>
<td>All habitats. Deposition rates above 25 kg N ha$^{-1}$ yr$^{-1}$.</td>
<td>Measure N concentrations in soil solution or surface water, calculate fluxes, and compare with observations for N-polluted systems e.g. (Roeve et al., 2006).</td>
<td>Pros: well-correlated with (higher) deposition rates; indicates advanced damage. Cons: unlikely to increase until later stages of N saturation</td>
</tr>
<tr>
<td>Metric</td>
<td>Type</td>
<td>Appropriate for</td>
<td>Recommended calculation method</td>
<td>Evaluation</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Mean ‘Ellenberg N’</td>
<td>M</td>
<td>Habitats where relationship with deposition has been demonstrated. All deposition rates.</td>
<td>Record plant species present, calculate mean Ellenberg N, and compare with typical values for the habitat e.g. using relationships from Stevens et al. (2011c).</td>
<td><strong>Pros:</strong> well-related to theoretical and observed effects of N on species-assemblages; can be modelled and also easily measured. <strong>Cons:</strong> Affected by factors other than N; meaning not immediately apparent.</td>
</tr>
<tr>
<td>Species richness</td>
<td>E</td>
<td>Grasslands, potentially other habitats such as mires. All deposition rates.</td>
<td>Record plant and lichen species present, calculate species richness, and compare with typical values for the habitat e.g. using relationships from Maskell et al. (2010).</td>
<td><strong>Pros:</strong> readily understood. <strong>Cons:</strong> affected by factors other than N; not applicable to all habitats.</td>
</tr>
<tr>
<td>Habitat Suitability Index (HSI)</td>
<td>E</td>
<td>All habitats. All deposition rates.</td>
<td>Mean simulated habitat suitability for ‘species of interest’ (Posch et al., 2014).</td>
<td><strong>Pros:</strong> potentially better-related to favourable conservation status than is species-richness. <strong>Cons:</strong> needs careful and transparent definition.</td>
</tr>
</tbody>
</table>

Dynamic models could also be adapted to simulate N dynamics in epiphytic / epilithic habitats.

The effects of N pollution on ecosystems are complex, and the temporal dynamics of impacts need to be considered. Although N pollution has some benefits for agricultural and forest productivity, untargeted applications of N are inefficient and have unintended consequences. The recommended metrics provide options for communicating and highlighting different aspects of N pollution, including pressure and impacts at different stages of ecosystem exposure. To develop management and policy responses it may sometimes be necessary to prioritise and/or combine the different metrics to make an overall assessment, although aggregate metrics can obscure genuine disagreements over the relative importance of different aspects of ecosystems (Suter, 1993). Reporting several distinct metrics has the advantage of separating pressure from response, and separating different aspects of response, and is useful for communicating the multiple facets of the N pollution problem.

**Acknowledgements**

The study was funded by the UK Department for the Environment, Food and Rural Affairs (DEFRA) under project AQ0823 (REBEND) and by the Centre for Ecology and Hydrology under project NEC05574. RJP is supported by the Russian Scientific Fund (Grant 14-14-00891). We are grateful to Max Posch of RIVM for permission to reproduce Figure 1a.

**References**


Coordination Centre for Effects, RIVM, NL.


de Vries, W., Posch, M., 2011. Modelling the impact of nitrogen deposition, climate change and nutrient limitations on tree carbon sequestration in Europe for the period 1900-2050.
Environmental Pollution 159, 2289-2299.


UNECE/Nordic Council of Ministers, Copenhagen, Denmark.


