TITLE: Nitrogen yield advantage from grass-legume mixtures is robust over a wide range of legume proportions and environmental conditions

AUTHORS: Suter, M., Connolly, J., Finn, J.A., Loges, R., Kirwan, L., Sebastià, M.T., Lüscher, A.

This article is provided by the author(s) and Teagasc T-Stór in accordance with publisher policies.

Please cite the published version.

The correct citation is available in the T-Stór record for this article.

NOTICE: This is the accepted version of the following article: Global Change Biology, doi: 10.1111/gcb.12880, Available online 28 January 2015. This has been published in final form at Wiley Online Library http://dx.doi.org/10.1111/gcb.12880. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document.

This item is made available to you under the Creative Commons Attribution-Non commercial-No Derivatives 3.0 License.
Nitrogen yield advantage from grass-legume mixtures is robust over a wide range of legume proportions and environmental conditions

Running head: Nitrogen yield from grass-legume mixtures

Suter, M.¹, Connolly, J.², Finn, J.A.³, Loges, R.⁴, Kirwan, L.⁵, Sebastià, M.T.⁶,⁷, Lüscher, A.¹

¹Agroscope, Institute for Sustainability Sciences ISS, Reckenholzstrasse, Zürich, Switzerland
²School of Mathematical Sciences, University College Dublin, Dublin 4, Ireland
³Teagasc, Environment Research Centre, Johnstown Castle, Wexford, Ireland
⁴Institut für Pflanzenbau und Pflanzenzüchtung, Christian-Albrechts-Universität, Kiel, Germany
⁵Waterford Institute of Technology, Cork Road, Waterford, Ireland
⁶Forest Sciences Centre of Catalonia, Solsona, Spain
⁷Dept HBJ, ETSEA, Universitat de Lleida, Lleida, Spain

Corresponding author: Matthias Suter, phone: +41 58 468 75 90, fax: +41 58 468 72 01, email: matthias.suter@agroscope.admin.ch

Keywords: sustainable agriculture, sustainable intensification, food security, N fertilizer replacement, climate change mitigation, climatic gradient, N uptake, protein, symbiotic N₂ fixation

Type of paper: Primary research article
Abstract

Current challenges to global food security require sustainable intensification of agriculture through initiatives that include more efficient use of nitrogen (N), increased protein self-sufficiency through home-grown crops, and reduced N losses to the environment. Such challenges were addressed in a continental-scale field experiment conducted over three years, in which the amount of total nitrogen yield \( (N_{\text{tot}}) \) and the gain of N yield in mixtures as compared to grass monocultures \( (N_{\text{gainmix}}) \) was quantified from four-species grass-legume stands with greatly varying legume proportions. Stands consisted of monocultures and mixtures of two \( N_2 \) fixing legumes and two non-fixing grasses.

The amount of \( N_{\text{tot}} \) of mixtures was significantly greater \( (P \leq 0.05) \) than that of grass monocultures at the majority of evaluated sites in all three years. \( N_{\text{tot}} \) and thus \( N_{\text{gainmix}} \) increased with increasing legume proportion up to one third of legumes. With higher legume percentages, \( N_{\text{tot}} \) and \( N_{\text{gainmix}} \) did not continue to increase. Thus, across sites and years, mixtures with one third proportion of legumes attained ~95% of the maximum \( N_{\text{tot}} \) acquired by any stand and had 57% higher \( N_{\text{tot}} \) than grass monocultures.

Realized legume proportion in stands and the relative N gain in mixture \( (N_{\text{gainmix}}/N_{\text{tot}} \) in mixture) were most severely impaired by minimum site temperature \( (R = 0.70, \ P = 0.003 \) for legume proportion; \( R = 0.64, \ P = 0.010 \) for \( N_{\text{gainmix}}/N_{\text{tot}} \) in mixture). Nevertheless, the relative N gain in mixture was not correlated to site productivity \( (P = 0.500) \), suggesting that, within climatic restrictions, balanced grass-legume mixtures can benefit from comparable relative gains in N yield across largely differing productivity levels.

We conclude that the use of grass-legume mixtures can substantially contribute to resource-efficient agricultural grassland systems over a wide range of productivity levels, implying important savings in N fertilizers and thus greenhouse gas emissions and a considerable potential for climate change mitigation.
Introduction

Global food security is currently challenged by the increasing demands for food, including meat and milk, which arise through the continuing growth of the world’s population and consumption (Godfray et al., 2010; Smith & Gregory, 2013). At the same time, food production is significantly affected by competition between food, feed and bioenergy, demands from other economic sectors for land and water, and the need to maintain and preserve ecosystem services and biodiversity (Thornton, 2010). Moreover, current food production is highly nitrogen (N) limited (Cassman et al., 2002), while the provision of industrial N is largely based on fossil energy and its multiple impacts on the environment (Galloway et al., 2008; Davidson, 2009; Canfield et al., 2010). Under a business-as-usual scenario, it must be assumed that any increase in food provision will further aggravate the pressure regarding the divergent demands for food security and environmental integrity (Tilman et al., 2002; Foley et al., 2005). This challenge has recently reinforced the need for more sustainable agriculture and sustainable intensification (Godfray et al., 2010; Foley et al., 2011).

Sustainable intensification of agriculture aims to raise productivity while at the same time reduce its environmental impacts (Godfray et al., 2010; Foley et al., 2011; Lüscher et al., 2014; Taube et al., 2014). In the face of the prevailing N limitation, rising costs of inorganic N fertilizers, and deleterious side-effects of excessive N application (Galloway et al., 2008; Canfield et al., 2010), increased sustainability and improved N self-sufficiency can be gained through home-grown N$_2$ fixing crops. Currently, grassland-based livestock production of medium to high management intensity depends largely on high-yielding pure grass stands requiring large inputs of mineral N fertilizers. Production and distribution of mineral N fertilizers need large amounts of energy (Kitani et al., 1999) and their application can result in
substantial N losses as nitrate (Ledgard et al., 2009) and greenhouse gases (GHG) to the environment (Schils et al., 2013; Schmeer et al., 2014). For example, each kg of N produced as ammonium nitrate in the industrial Haber-Bosch process consumes 58 MJ of energy and emits 8.6 kg CO$_2$ equivalents (Kitani et al., 1999; Ecoinvent Centre, 2010). Moreover, according to the guidelines of the IPCC (2006), for every 100 kg of N fertilizer added to the soil, on average 1.0 kg of N is emitted as N$_2$O, a GHG that is approximately 300 times more potent than CO$_2$. At the same time, European livestock systems depend strongly on protein imported from overseas. A major challenge, therefore, is to increase home-grown forage protein with reduced input of mineral N fertilizers and at the same time reduce N losses to the environment (Peyraud et al., 2009; Lüscher et al., 2014; Taube et al., 2014).

Grass-legume mixtures offer the benefit of symbiotic N$_2$ fixation by legumes, which are able to utilize atmospheric N$_2$ for their requirements and thereby produce more protein with less N input. The amount of symbiotic N$_2$ fixation by legumes can be substantial and ranges from 100 to 380 kg ha$^{-1}$ year$^{-1}$ in northern temperate/boreal regions (Ledgard & Steele, 1992; Carlsson & Huss-Danell, 2003). Grown in mixtures with grasses, legumes meet their own N demand by deriving more than 80% from symbiosis (Boller & Nösberger, 1987; Zanetti et al., 1997; Nyfeler et al., 2011) and consequently, the relative availability of soil N increases for grasses (“N sparing”, Temperton et al., 2007). Besides symbiotic N$_2$ fixation, other processes have been found to increase yield and efficiency in resource uptake by grass-legume mixtures. These include facilitation, that is, N transfer from legumes to grasses (Høgh-Jensen & Schjoerring, 1997; Pirhofer-Walzl et al., 2012; Rasmussen et al., 2013) and increased exploitation of soil resources through spatial (deep- and shallow-rooting) or temporal niche complementarity in resource uptake (van Ruijven & Berendse, 2005; Mueller et al., 2013). All of these processes can lead to considerable gains in N yield of mixtures compared to grass
monocultures; consequently, the use of such mixtures in agricultural grassland systems could allow substantial reductions in the application of industrial N fertilizers.

Inclusion of legumes in grassland management has been proposed as an important strategy for climate change mitigation in the agricultural sector (Smith et al., 2008; Smith & Gregory, 2013; Bustamante et al., 2014). The use of grass-legume mixtures in temporary grassland is widely applicable and does not necessarily compromise harvest yield, thereby improving agricultural efficiency (Kirwan et al., 2007; Finn et al., 2013). Moreover, the application is practical and robust, as demonstrated over many years by the “Swiss Standard Mixtures” that use up to eight species of grasses and legumes in temporary grassland to improve legume persistence under various pedo-climatic conditions (Frey, 1955; Suter et al., 2012). Grass-legume mixtures offer also a sustainable farm management practice: there is no evidence of significant emissions of N$_2$O arising from the process of symbiotic N$_2$ fixation (Rochette & Janzen, 2005; Barton et al., 2011) and thus, N$_2$ fixation has been removed as a direct source of N$_2$O in the revised GHG guidelines of the IPCC (2006). Finally, mixing grasses with legumes is a mitigation measure that can be implemented in the near future, which is an important feature for meeting emission reduction targets (Smith et al., 2013; Bustamante et al., 2014).

Although amounts of total N yield and N from symbiotic sources have been quantified in temperate grassland (Ledgard & Steele, 1992; Carlsson & Huss-Danell, 2003), data from arctic or continental ecosystems are rare. In such environments, low winter temperatures and/or precipitation may hamper the legumes’ growth, and accordingly, the benefit of including legumes in mixtures to achieve high gains in N yield may be limited. Absolute amounts of N from symbiosis seem to decrease at higher latitudes of Europe (Nesheim &
Oyen, 1994); however, there is no simple correlation between latitude and the N performance of legumes (Carlsson & Huss-Danell, 2003). Any large-scale geographic influence acts in concert with local pedo-climatic conditions that also interact with the type and intensity of management, all of which may strongly affect the legumes’ growth.

The COST Action 852 entitled “Quality legume-based forage systems for contrasting environments” (www.cost.eu/domains_actions/fa/Actions/852) aimed at increasing the quantity and quality of home-grown protein from regionally adapted legume-based forage systems. To this aim, a coordinated continental-scale field experiment was established, the “Agrodiversity Experiment”, focusing on intensively managed, mown temporary grassland as a model system. Monocultures and mixtures comprised of two grasses (non N\textsubscript{2} fixing) and two legumes (N\textsubscript{2} fixing), and were set up to result in a broad range of legume proportion in stands (Kirwan et al., 2007). Here, we report on the benefits of legumes to total N production in these grass, legume, and mixed swards by analyzing the total N yield (\textit{N}_{\text{tot}}) and realized legume proportion of swards, and the potential N yield gain in mixtures as compared to grass monocultures (\textit{N}_{\text{gainmix}}). To our knowledge, this is the first study to directly relate N yields in mixtures and monocultures to greatly differing stand legume proportions across largely contrasting pedo-climatic conditions, which will reveal the potential of these grassland systems for resource-efficient, sustainable agriculture through savings in N fertilizers and thus GHG emissions. The following specific questions were addressed:

1. Is \textit{N}_{\text{tot}} consistently higher in grass-legume mixtures as compared to grass monocultures?

2. Is \textit{N}_{\text{tot}} and \textit{N}_{\text{gainmix}} affected by legume proportion in the sward? If yes, how much does the effect of legume proportion on \textit{N}_{\text{tot}} and \textit{N}_{\text{gainmix}} vary across sites?

3. Is the effect of legume proportion on \textit{N}_{\text{tot}} and \textit{N}_{\text{gainmix}} persistent over three years? Does legume proportion itself persist over time?
4. Are realized legume proportion, \( N_{\text{tot}} \) in mixture, and the relative N gain in mixture 
\( (N_{\text{gainmix}}/N_{\text{tot}} \) in mixture) related to climatic variables and to productivity levels across sites?
Material and Methods

Experimental design

At each of sixteen sites and following a common protocol, four monocultures and eleven mixtures were established at two levels of seed density for a total of 30 stands per site. The sites spanned a gradient of climate from Atlantic to continental and from temperate to arctic (Fig. 1, Table S1, Supporting Information). Regarding climate and productivity, mean annual temperature of the sixteen sites ranged between 4.2 and 10.9 °C, precipitation between 492 and 1556 mm year\(^{-1}\), and productivity of aboveground biomass between 2 and 14 t DM ha\(^{-1}\) year\(^{-1}\) (see below). Full information to all sites is given in Kirwan et al. (2014), where geographic coordinates and further details on climate and soils are provided. The selection of species used for experimentation focused on i) species known to achieve high forage quantity and quality in systems of intensive grassland management, and ii) functional types of species that were expected to maximise complementarity in resource use. Four functional types of species were initially defined as the factorial combination of traits being associated with the manner of N acquisition (non N\(_2\) fixing grasses vs. N\(_2\) fixing legumes), and temporal pattern of species’ growth (fast establishment vs. temporally persistent), resulting in the following set: fast establishing grass (G1), temporally persistent grass (G2), fast establishing legume (L1), temporally persistent legume (L2). These four functional types were consistent across all sixteen sites although the species selected at sites varied depending on geographical region (Table 1).

The four monocultures consisted of one of each of G1, G2, L1, or L2, and mixtures contained all four functional types of species in varying relative abundances following a simplex design (Cornell, 2002). This was achieved by systematically varying the sown species proportions in mixtures (Table S2, Supporting Information) to result in four mixtures dominated in turn by
one species (70% of one species, 10% of the three others), six mixtures co-dominated in turn
by pairs of species (40% of each of two species, 10% of the two others), and one mixture with
equal species proportions (25% of each species). Doing so, sown legume proportions across
different communities were 0, 20, 50, 80, and 100% (see also Kirwan et al., 2007 for further
details on the design).

184 *Maintenance and measurements*

185 Plots were fertilized with commercial N fertilizer. The amount was constant for all plots and
for all years of experimental duration within individual sites, but varied among sites to range
from 0 to 150 kg N ha\(^{-1}\) year\(^{-1}\) (Table S1, Supporting Information). This range of N
application reflected background productivity levels across the large climatic gradient as well
as variations in types and fertility of soils. Note that the maximum level of 150 kg N ha\(^{-1}\) year\(^{-1}\)
is far below highest levels of N application to grass monocultures in intensive grassland
management and that N application of up to 150 kg ha\(^{-1}\) year\(^{-1}\) did not impair positive grass-
legume interactions at one of the study sites (Nyfeler et al., 2009). Background levels of
phosphorus and potassium in experimental plots were adjusted to non-limiting amounts
(Kirwan et al., 2014). Aboveground biomass of plots was harvested several times per year
following the agronomic practice at each site (Table S1, Supporting Information), and
representative subsamples of harvested yield were sorted into the four sown and pooled
unsown species. Drying to constant weight and summing over harvests allowed computing of
the total harvested dry mass per year and species’ fractional contributions. Importantly,
measurments were only recorded in the first three full years of production in this temporary
grassland; thus, the sowing year was not considered. This restriction was imposed to evaluate
the fully established system.
Record of climatic data

Data on climatic conditions were monitored by weather stations at each of the sixteen sites. Daily precipitation was summed and mean daily temperature was averaged across all days of the year to give annual precipitation and mean annual temperature per site. Moreover, daily minimum and maximum temperatures were used to compute the annual minima and maxima per site as the average of the ten days with most extreme values. This was of specific interest because extreme temperatures are assumed to especially impair legumes’ growth and N\textsubscript{2} fixation (Lynch & Smith, 1993; Zhang \textit{et al.}, 1995).

Analysis of N concentration in plant material

Nitrogen concentration (N\textsubscript{conc}) was measured at each harvest from a representative subsample of total harvested biomass from each plot, with biomass being dried to constant weight at 65 °C and ground to pass through a 1 mm sieve. The value of N\textsubscript{conc} of all samples was determined by near-infrared reflectance spectroscopy (NIRS) at one lab (Christian-Albrechts University, Kiel, Germany) using a NIRSystems 5000 monochromator (FOSS, Silver Spring, USA). See Appendix A, Supporting Information, for validation of the NIRS method. For a subset of sites, N\textsubscript{conc} was not measured at all three experimental years and from all plots of the design (see Table S1, Supporting Information). Note that temporary grassland in crop rotation systems is generally maintained for a sowing year and one or two subsequent production years. Here, we measured N\textsubscript{conc} for two production years for most sites (year 1 and 2), while half the sites also had data for a third year (year 3), resulting in a total of 350, 304, and 167 analyzed plots in years 1, 2, and 3, respectively. Importantly, included sites covered the full pedo-climatic range in all years. There were a total of 36 site-years of data.
227 Data analyses

228 We aimed to analyze the total annual N yield \((N_{tot})\) per plot. To achieve this, values of \(N_{conc}\) of each harvest were first multiplied by total dry matter per harvest to receive the N yield per harvest and plot. Annual \(N_{tot}\) per plot was then computed as the sum over all harvests and reflects therefore an N output of the system that is seasonally weighted for variability in \(N_{conc}\) and biomass yield.

233

i) Comparisons of mixtures against grass monocultures

234 Annual \(N_{tot}\) of mixtures was initially compared against \(N_{tot}\) of the average grass monoculture, as this comparison reflected a test of mixed stands (all of which included legumes) against stands sown only with grass species. Wilcoxon rank sum tests were applied for inference.

ii) Testing the effect of legume proportion on total nitrogen yield

239 The effect of legume proportion \((P_{leg})\) the on annual \(N_{tot}\) was analyzed applying a regression approach following Nyfeler et al. (2011), which used the following basic notation:

\[
N_{tot} = f_0 + p_1 P_{leg} + p_2 P_{leg}^2 + p_3 P_{leg}^3 + aD + yDeltBiomass + E \quad \text{eqn. 1}
\]

243 In this formulation, \(P_{leg}\) denotes the summed proportions of the two legumes L1 and L2, using the harvested biomass proportions of the previous year (sown proportions were used for year 1) as predictor variables to overcome confounding effects of year-to-year changes in community composition on \(N_{tot}\). Thus, \(N_i\) to \(N_6\) estimate the linear and non-linear effects of stand legume proportion on \(N_{tot}\). The intercept, \(P_0\), gives the estimate of \(N_{tot}\) if \(P_{leg} = 0\), i.e. for grass monocultures including potential weeds. The effect of seed density is estimated by \(a\), with \(D\) denoting the level of seed density coded as -1 and +1 for low and high density, respectively, so all other terms are estimated at average density. The effect of fluctuations in stand biomass \((DeltBiomass)\) on the response variable is estimated by the \(y\) coefficient (for
DeltaBiomass see Appendix A, Supporting Information). The error $E$ is assumed normally distributed with zero mean and variance $a^2$.

Equation 1 was extended to a linear mixed model where all coefficients were estimated for each of three years, and where random coefficients were added to estimate the general stand performance while allowing for variation across sites for each of the three years (Pinheiro & Bates, 2009) (see eqn. S1, Appendix A, Supporting Information, for detailed notation).

Inference on fixed main effects of the linear mixed model was based on single term deletion from a main effects model as specified in eqn. 1 (each effect in turn, including $\sim_{eg}^2$ and $\sim_{eg}^3$) and subsequent $F$-tests for comparison of models while applying the Kenward-Roger method to determine the approximate denominator degrees of freedom (Kenward & Roger, 1997); interactions were similarly tested but from a model that included all effects. The range of legume proportion for which $N_{tot}$ was significantly different from its maximum was computed using the Johnson Neyman technique (Johnson & Neyman, 1936) as applied in Suter et al. (2007).

In regressions of eqs. 1 and S1 (Appendix A, Supporting Information), species G1, G2, L1, and L2 are grouped into grasses and legumes according to their functional ability to fix $N_2$ (or not), and such analysis does not include interactions between species regarding temporal pattern of species’ growth. This grouping was justifiable because preliminary analyses revealed that interactions affecting $N_{tot}$ were mainly between non-fixing grasses and $N_2$ fixing legumes, but to a far lesser extent between fast establishing and temporally persistent species within grasses and legumes (see Appendix A, Supporting Information, for details and inference on pooling individual species performances).
iii) Computing gain of N yield in mixture

Total N yield in mixed grass-legume stands and the respective monocultures can be subdivided into different parts to illustrate the role of legumes and their interactions with grasses. In grass monocultures (legume proportion = 0), $N_{\text{tot}}$ accumulated by the sward derives solely from the soil and from fertilizer N (quantity A, dashed white line for reference, Fig. 2). In legume monocultures (legume proportion = 1), $N_{\text{tot}}$ is often greater (A + B), due mainly to symbiotic N$_2$ fixation of legumes. Mixing grasses with legumes should lead therefore at first to a linear increase of accumulated N by the stand with increasing mixture legume proportion through the contribution of the legume component, i.e. through a constant amount of N added per unit of legumes (continuous white line). In addition, positive grass-legume interactions such as stimulation of the rate of symbiotic N$_2$ fixation (% N derived from symbiosis in the legume) when grass is present (Nyfeler et al., 2011), N transfer (Høgh-Jensen & Schjoerring, 1997; Zanetti et al., 1997) or increased utilization of fertilizer and soil N resources through temporal and/or spatial niche complementarity (Mueller et al., 2013) can result in a nonlinear surplus of N yield (C, Fig. 2), resulting in a total N yield of A+B+C in mixed swards.

This study aimed to quantify the amount of N yield gain in mixtures as compared to grass monocultures. This was achieved by computing the difference between $N_{\text{tot}}$ in mixture and $N_{\text{tot}}$ in grass monoculture using the estimates of the regression model (based on eqn. S1, Appendix A, Supporting Information). This quantity of N yield gain in mixture is hereafter termed $N_{\text{gainmix}}$ ($= N_{\text{tot}}$ in mixture minus $N_{\text{tot}}$ in grass monoculture, quantity B + C in Fig. 2), and was calculated both for the mean across all sites (using the fixed parameter estimates of the regression) and for each of the sixteen individual sites (using the variation around the fixed mean). Note that the quantity of $N_{\text{gainmix}}$ reflects the total of N gain in mixture compared...
to grass monoculture, but it does not allow quantification of the individual contributions from
symbiotic N\textsubscript{2} fixation, N transfer, more efficient exploitation of soil and fertilizer N, or N
from any other source (e.g. decaying roots of legumes).

iv) Relating N data to productivity and climate
To evaluate patterns in N dynamics over the environmental gradient, realized legume
proportion, N\textsubscript{tot} in mixture, and the relative N yield gain in mixture (N\textsubscript{gainmix}/N\textsubscript{tot} in mixture)
were related to site productivity, annual precipitation, mean annual temperature, and
minimum and maximum temperature. To increase the robustness of results, these calculations
were based on site means across the first two experimental years. For consistent comparison,
both N\textsubscript{tot} in mixture and the relative N gain were calculated for a mixture with one third
proportion of legumes and two thirds of grasses using a simplified linear mixed model (eqn.
S2, Appendix A, Supporting Information). The ratio of proportions (1/3:2/3 legumes:grasses)
could be justified from analyses that regressed N\textsubscript{tot} on legume proportion (see results), and
amounts of total N yield for this representative mixture are hereafter termed N\textsubscript{totmix}. Site
productivity was estimated by averaging across the biomass yields of all grass monocultures.
Pearson correlation and ordinary least squares regression were used to quantify relationships
between legume proportion, N\textsubscript{totmix}, the relative N yield gain in mixture (N\textsubscript{gainmix}/N\textsubscript{totmix}), and
environmental variables (predictors) (see Table S4, Supporting Information, for site values of
the five environmental variables, and Table S5 for their correlation matrix). All analyses were
performed with the statistics software R (R Development Core Team, 2014).
Results

Positive effect of grass-legume mixtures on total N yield

Annual $N_{\text{tot}}$ of mixtures was significantly greater than that of grass monocultures at the majority of evaluated sites in all years (Fig. 3), and approximated the values of $N_{\text{tot}}$ in legume monocultures. In general, the positive mixture effect on $N_{\text{tot}}$ was already evident in the first year and persisted over years (Fig. 3); however, no mixture effect on $N_{\text{tot}}$ was apparent at the continental sites Lithuania_a (site 18; years 1 and 2), and Poland_a and Poland_b (sites 26 and 27, all years).

Positive effect of legume proportion on total N yield and N yield gain in mixture

$N_{\text{tot}}$ was significantly affected by the proportion of legumes in mixtures. The effects of legume proportion on $N_{\text{tot}}$ were similar across years and were maximal when legume proportions in mixtures were 60, 42, and 40% in years 1, 2, and 3, respectively (Fig. 4). More importantly, because $N_{\text{tot}}$ was non-linearly affected by legume proportion (Table 2, Table S6, Supporting Information), $N_{\text{tot}}$ was not significantly different from the maximum amount in mixture over a wide range of legume proportion (Fig. 4), and mixtures with approximately one third proportion of legumes attained 95% of maximum $N_{\text{tot}}$ (Table 3). $N_{\text{tot}}$ varied substantially among sites, with greatest (predicted) values in mixtures being as high as 480 kg N ha$^{-1}$ year$^{-1}$ at Ireland (site 15, year 1) and Switzerland (site 34, year 2), and smallest values in Iceland (sites 13 and 14: around 50 kg N ha$^{-1}$ year$^{-1}$; Fig. 4).

Because mixtures with approximately one third proportion of legumes attained 95% of the maximum $N_{\text{tot}}$ (Table 3), further values for $N_{\text{tot}}$ and $N_{\text{gainmix}}$, below were estimated for this representative mixture (1/3:2/3 legumes:grasses). Doing so, $N_{\text{totmix}}$ decreased by 12% between
years 1 and 2 ($P = 0.09$ for the difference in $N_{\text{totmix}}$), and by a further 23% between years 2 and 3 ($P = 0.002$).

Average annual values of $N_{\text{gainmix}}$ across sites ranged between 108 and 76 kg ha$^{-1}$ year$^{-1}$ (years 1 and 3, respectively), resulting in a relative N gain in mixture ($N_{\text{gainmix}}/N_{\text{totmix}}$) that exceeded 0.3 in all years (Table 3). Again, there was a large variation of $N_{\text{gainmix}}$ among sites with amounts being as high as 280 kg ha$^{-1}$ year$^{-1}$ at Norway_a (site 22, year 1) but being close to zero at the continental sites Lithuania_a (site 18, year 1), Poland_a, and Poland_b (sites 26 and 27, all years), suggesting that at these sites legumes in mixtures induced no gain in total N yield (Fig. 4). Compared to grass pure stands and averaged across sites, $N_{\text{gainmix}}$ was 61, 46, and 64% in years 1, 2, and 3, respectively (calculated from Table 3), meaning that mixtures with only one third proportion of legumes had, on average across all years, 57% higher total N yield than pure grass stands.

Declining stand legume proportion over time

Realized legume proportion in stands decreased generally over time: across sites, realized legume proportion was 32, 21, and 9% in years 1, 2, and 3, respectively. Regarding site variation, almost the full possible range was covered: over all swards with highly varying legume proportion in the seed mixture, Norway_a (site 22), Wales_a (35), and Switzerland (34) still had average legume proportions around 50% and maximal values exceeding 80% by year 2; in contrast, legumes were absent from the second year onward in Ireland (15) and had disappeared by the third year in Lithuania_c (20) (Fig. 4).
Strong effect of minimum site temperature on legume proportion and N yield gain in mixture

Out of the environmental variables tested, minimum site temperature most strongly impaired both stand legume proportion and the relative N yield gain in mixture (Fig. 5a & c), indicating that harsh environmental conditions hampered legumes’ growth (see Table S4, Supporting Information, for site values of $N_{\text{totmix}}$ and $N_{\text{gainmix}}$, and Table S7 for a summary of all regressions). Although to a weaker degree, low annual precipitation also impacted on both stand legume proportion and $N_{\text{gainmix}}/N_{\text{totmix}}$ (Fig. 5b & d). The parallel responses of legumes’ growth and relative N yield gain in mixture to these environmental parameters suggested that legume proportion and $N_{\text{gainmix}}/N_{\text{totmix}}$ were correlated to each other, which indeed was the case (Fig. 5e). Furthermore, $N_{\text{totmix}}$ was positively correlated to site productivity ($R = 0.703$, $P = 0.003$, Table S7, Supporting Information); however, the relative N yield gain in mixture was not (Fig. 5f). Taken together, this suggested that the relative N gain of a balanced grass-legume mixture was affected more by climatic conditions than by the largely differing productivity levels across sites.
Discussion

Mixing grasses and legumes in agricultural grassland systems yielded considerable benefits to total N yield ($N_{\text{tot}}$) over a wide range of European environmental conditions. Averaged across sites and years, mixtures with only one third proportion of legumes had as much as 57% higher $N_{\text{tot}}$ than pure grass stands, indicating a substantial N yield gain in mixtures ($N_{\text{gainmix}}$), and such mixtures attained amounts of $N_{\text{tot}}$ close to the maximum $N_{\text{tot}}$ acquired by any stand.

The relative N yield gain in mixture ($N_{\text{gainmix}}/N_{\text{totmix}}$) was not affected by the great differences in productivity levels across sites. This clearly highlights the potential of grass-legume mixtures as a practical management option for sustainable agriculture, and we elaborate on a number of relevant points below.

Substantial N yield advantage in mixtures as compared to grass monocultures

Amounts of $N_{\text{tot}}$ from the temperate sites of this study corresponded well to reports from comparable grass-legume systems (Boller & Nösberger, 1987; Høgh-Jensen & Schjoerring, 1997; Carlsson & Huss-Danell, 2003 for review); however, this experiment also covered pedo-climatic regions that have been poorly investigated so far. At the majority of sites, the positive mixture effect on $N_{\text{tot}}$ appeared in the first year, was maintained for the second year and was still apparent in five out of eight evaluated sites in the third year. In particular, the positive mixture effect on $N_{\text{tot}}$ was apparent also at the Nordic sites in Iceland (sites 13 and 14) and Tromsø-Norway (site 23), where absolute amounts of $N_{\text{tot}}$ (Fig. 3) and harvested biomass yield (Finn et al., 2013) were comparably small. Highest N advantages of mixtures versus grass monocultures were more than 250 kg N ha$^{-1}$ year$^{-1}$ (Belgium (site 1), Switzerland (site 34), Fig. 3a) and resulted in a more than two-fold higher N output in mixed swards than in stands with grass only. Considering the substantial differences in site productivity, ranging between 2 and 14 t DM ha$^{-1}$ year$^{-1}$, and the substantial gain of $N_{\text{tot}}$ in mixtures across sites and
years, the data clearly demonstrate the great potential of mixed grass-legume swards for resource-efficient production in varying pedo-climatic conditions. Using mixed swards instead of pure grass stands, more N yield can be expected for a given amount of N fertilizer applied. Alternatively, if the aim is to reduce fertilizer N application for financial, regulatory and/or environmental reasons (Godfray et al., 2010), our data show the potential to do so without necessarily compromising N yield and total harvested biomass (Nyfeler et al., 2009; Finn et al., 2013).

416

417 N yield gain in mixtures is evident over a wide range of legume proportion

418 With few exceptions, \( N_{\text{tot}} \) demonstrated a considerable increase with increasing legume proportion up to about 30%. With higher percentages of legumes, \( N_{\text{tot}} \) did not continue to increase (Fig. 4). This is a highly valuable result. It indicates that almost all (~95%) of the maximum benefit to \( N_{\text{tot}} \) from mixing grasses and legumes can be achieved with a modest (~30%) legume proportion in the mixture. Moreover and equally relevant, the wide range for which \( N_{\text{tot}} \) was not different from maximum values indicates that the benefits of legumes regarding N gain in mixtures can prevail despite considerable fluctuations of grass-legume proportions in swards. This is an important feature for practical grassland management.

426

427 The method to compute \( N_{\text{gainmax}} \) has been previously used to calculate the amount of apparent symbiotic N\(_2\) fixation (\( N_{\text{sym}} \)) in grass-legume mixtures (“N-difference method“, e.g. Ledgard & Steele, 1992). The N-difference method has been criticized because it assumes that the cumulative uptake of N from soil and fertilizer of grasses and legumes in mixture would be the same as for the reference grass monoculture, which may not necessarily be the case (Boller & Nösberger, 1987; Nyfeler et al., 2011). The N-difference method therefore can lead to biased estimates of \( N_{\text{sym}} \). At the Swiss site of our study, Nyfeler et al. (2011) used the more
accurate $^{15}$N dilution procedure to quantify $N_{\text{tot}}$, $N_{\text{sym}}$ and N uptake from non-symbiotic sources (N derived from the soil and/or fertilizer). They demonstrated that, for a 50:50 mixture of grasses and legumes and fertilizer N application of 50 kg ha$^{-1}$ year$^{-1}$, the N-difference calculation resulted in $N_{\text{sym}}$ values that were 26% higher than the $^{15}$N dilution procedure in the first production year, while overestimation was 24% in the second year. For N applications of 150 kg ha$^{-1}$ year$^{-1}$, overestimation by the N-difference method became smaller, and was 17% and 7% in the first and second year, respectively. In Northern mixed grassland (Vågønes, Norway, 69°N) of *Trifolium repens* L. (white clover) and *Phleum pratense* L. (timothy) and no fertilizer N application, the overestimation of the N-difference method compared to $^{15}$N dilution was 7% and 36% in two production years, respectively (Nesheim & Oyen, 1994). Regarding our continental-scale experiment, $N_{\text{gainmix}}$ corresponds to values of the N-difference calculation. It is important to realize that this difference method is appropriate to the purpose used here, namely to measure the total N gain achieved by the grass-legume mixture compared to the pure grass stand. However, because it appears that this difference method does not allow quantification of symbiotic N$_2$ fixation and N uptake from non-symbiotic sources with acceptable accuracy, we do not aim to specify these different fractions. Nevertheless, the above assessment suggests that the great majority of $N_{\text{gainmix}}$ is derived from N$_2$ fixation of legumes (see also Boller & Nösberger, 1987; Zanetti et al., 1997).

This gain of N yield in mixtures due to the presence of legumes can be explained by various mechanisms. A linear increase in $N_{\text{tot}}$ with increasing legume proportion can be expected solely through a constant additional N input to the system per unit of legumes (see B in Fig. 2). However, because we found a highly nonlinear response of $N_{\text{tot}}$ (Fig. 4, C in Fig. 2), positive interactions between grasses and legumes must have played a role, and we suggest four possible mechanisms. First, positive mixing effects have been attributed to stimulation of
the rate of symbiotic N$_2$ fixation by N-demand from co-occurring grasses (Høgh-Jensen & Schjoerring, 1997). At the Swiss site, Nyfeler et al. (2011) revealed that the rate (%) of N derived from symbiosis in the clover plants was significantly higher at low legume proportion than at high legume proportion. This pattern could be explained by strong competition from grasses for N from non-symbiotic sources: even if the grass component in mixture was reduced to 50%, grass still took up the same total amount of N from non-symbiotic sources as a stand with 100% grass (Nyfeler et al., 2011). Second, the presence of both grass and legume components in mixture allows for transfer of symbiotically fixed N from legumes to grasses. Quantification of apparent N transfer between legumes and grasses shows that more than 40% of the N nutrition of the grass component in mixture can derive from N$_2$ fixation of the legumes (Boller & Nösberger, 1987; Høgh-Jensen & Schjoerring, 1997; Nyfeler et al., 2011) with N transfer being greatest in equilibrated mixtures. Third, the nonlinear response of N$_{tot}$ can be attributed to an increased capability of the grass component to acquire N from non-symbiotic sources in the presence of legumes (Nyfeler et al., 2011). Grasses have a denser root system (i.e. comparably higher root length density per unit soil volume and high root surface area of active absorption, Craine et al., 2002; Hill et al., 2006), which they use to outcompete the legumes in accessing the available mineral and fertilizer N. Fourth, increased utilization of fertilizer and soil N resources through temporal and/or spatial niche complementarity between grasses and legumes can also explain the positive mixing effects (van Ruijven & Berendse, 2005; Mueller et al., 2013). To conclude, the benefit of mixed swards to total N output of the system comes through significant grass-legume interactions.

Considerable mixing effects on total harvested biomass yield have recently been shown for the same sites as presented here (Finn et al., 2013). Finn et al. (2013) found transgressive overyielding (mixtures outperformed highest yielding monocultures, Trenbath, 1974) of about
20% (excluding weed biomass), indicating that mixtures produced more biomass than either
grass or legume monocultures. To maximise both total N output and total biomass yield,
while at the same time minimising the risk of N losses to the environment (Loiseau et al.,
2001; Nyfeler, 2009), we infer that an optimal range of legume proportion in mixtures should
range between 30-50%.

Legume proportion and N yield gain in mixtures are affected by environmental conditions
We found evidence that legume proportion and the relative N yield gain in mixture were
negatively correlated with minimum winter temperatures and low annual precipitation (Fig.
5). Much of the evidence on restrictions of legumes’ growth and symbiotic N\(_2\) fixation comes
from studies in growth chambers (e.g. Nesheim & Boller, 1991; Lynch & Smith, 1993; Zhang
et al., 1995; Serraj & Sinclair, 1996). Low temperature seems to hamper N\(_2\) fixation more
than plant growth (Lynch & Smith, 1993; Zhang et al., 1995; Hartwig, 1998), and severe
drought has been shown to disrupt nodule activity (Serraj & Sinclair, 1996; Serraj et al.,
1999). Such results accord with our findings gained under field conditions. However, because
we did not directly measure symbiotic activity of legumes but demonstrated impacts of
climate on legume proportion, we cannot conclude whether harsh climatic conditions
impacted more on symbiotic N\(_2\) fixation or on legume plant growth.

A decrease in the legume proportion of mixed swards can be induced by application of N
fertilizers. For example, in a related experiment (Nyfeler et al., 2009), different levels of
mineral N were applied to grass-clover mixtures (two grass and two clover species) managed
for three consecutive years. Here, N fertilization significantly affected clover proportion:
averaged across years and all mixtures receiving 50 kg N ha\(^{-1}\) year\(^{-1}\), the two clover species
(sum of both) achieved 41% proportional biomass; however, fertilized with 150 kg N ha\(^{-1}\)
Clover proportions were only 31% (Nyfeler et al., 2009). Comparable results were found by Nassiri & Elgersma (2002), where in only one growing season the application of 150 kg N ha\(^{-1}\) year\(^{-1}\) reduced the clover content in dry matter harvest of perennial ryegrass-white clover mixtures to 12% compared to 43% without N application. Such impacts of N fertilizers on legume persistence might be of less importance in temporary grassland as part of crop rotation systems, where swards are generally maintained for a seeding year and one or two production years before they are ploughed for growing cereal crops. Moreover, although legumes might decrease over time, their positive effect on mixture biomass yield can still be prevalent even at low proportion (Nyfeler et al., 2009; Finn et al., 2013), which can be explained by strong grass-legume interactions as outlined above or by legacy effects (e.g. release of fixed N from decaying roots or increased N pools in soil organic matter derived from N\(_2\) fixation of preceding years).

In our experiment, positive effects of legumes on N\(_{\text{tot}}\) were still evident in year two and three (Fig. 3) despite the successive decrease in legume proportion, and mixtures with only one third proportion of legumes provided a significant gain in N yield as compared to pure grass stands (Fig. 4). This proves our experiment with sown temporary grassland to be a good model system to study the relation between legume proportions realized in the sward and N yield gain in mixtures, and this relation holds also for permanent grassland as long as adequate proportions of legumes can be maintained. Yet, the persistence of legumes in permanent grassland is challenging (Guckert & Hay, 2001), and our study implies that research should focus on sward management strategies to stabilize legume proportion under varying climatic conditions. Evidence suggests that adjusting N fertilizer rates and defoliation frequency can increase the abundance of white clover (Schwank et al., 1986; Hebeisen et al., 1997; Lüscher et al., 2014), the most important legume species in permanent grassland of
temperate regions (Guckert & Hay, 2001). In our experiment, no specific actions were undertaken to counteract legume decrease. Also, realized legume proportion was not correlated to N application rates at sites ($R = -0.219, P = 0.434$), indicating that environmental conditions were more decisive for legume performance than N fertilizers at rates applied in our experiment.

Four sites merit further consideration. At the three continental sites Lithuania_a, Poland_a, and Poland_b, legumes established well in the first year and reached proportions of up to 96%; nevertheless, N gains in mixture were not evident (Fig. 4a) suggesting that symbiotic N$_2$ fixation of the clovers did not work properly. We suspect that in concert with low winter temperatures hampering nodulation (Hartwig, 1998), lack of $Rhizobia$ species in soils of the agronomically improved grassland may have prevented legumes to form an efficient symbiotic relationship. The situation is different for Ireland, where legumes were present and mixtures realized high gains of total N yield in the first year, but legumes disappeared from the second year onward. With minimum temperatures of -1° C, annual precipitation around the mean of the investigated range (932 mm year$^{-1}$), and comparably high site productivity (14.1 t DM ha$^{-1}$ year$^{-1}$), climatic conditions cannot be responsible for the strong legume decline. Because the Irish site had been a highly fertilized monoculture of $L. perenne$ grassland for many years preceding the experiment, we speculate that the very high soil fertility and resulting high competition from the grasses during the experiment may have hampered the legumes’ growth (Schwank et al., 1986). Also, Ireland had a soil pH of 5.3, being the lowest among sites (Kirwan et al., 2014), which may have hampered the growth of the two legume species.
Taken together, it can be concluded that climatic conditions and the natural presence of $\text{Rhizobia}$ species define a boundary for the performance of legumes, in particular for the species used in this experiment. Our results indicate the need for specific research on how legumes can be maintained in mixed swards under varying and extreme climatic conditions.

Besides the inoculation of soil with $\text{Rhizobia}$ species and the optimization of sward management through adjusted N fertilizer inputs and defoliation frequencies, research should focus on breeding of adapted cultivars and selection of different legume species that would withstand low winter temperatures and/or severe drought.

**Wider implications: Legumes as a key contributor to sustainable intensification of grassland across largely differing productivity levels**

One important result of this study is that, although $N_{\text{tot}}$ in mixture was strongly affected by site productivity, the relative N yield gain in mixture was not (Fig. 5f). This means that less productive sites as well as more productive sites can equally profit from grass-legume mixtures to increase N output. Adaptation of legumes to differing productivity levels can be explained by the concept of “N$_2$ fixation regulation by demand” (Hartwig, 1998; Soussana & Tallec, 2010). Following Hartwig (1998), the degree of symbiotic N$_2$ fixation of legumes is controlled by a series of eco-physiological triggers and N feedback mechanisms from the individual plant to the ecosystem level, with N$_2$ fixation of legumes being largely regulated by the N sink strength (N-demand) of the whole system (Hartwig, 1998; Soussana & Tallec, 2010; Lüscher et al., 2011). It has been shown repeatedly that, under low to medium N fertilizer supply (< 100 kg N ha$^{-1}$ year$^{-1}$), legumes in a balanced mixture with grasses acquire the large majority of their N nutrition through symbiotic N$_2$ fixation (Nesheim & Oyen, 1994; Høgh-Jensen & Schjoerring, 1997; Nyfeler et al., 2011). In contrast, as amounts of N fertilizer increase, a decline of N acquired from symbiosis in legumes has been demonstrated.
studies and mathematical models (Soussana et al., 2002; Soussana & Tallec, 2010). Such propensity to buffer N supply can also be an advantage for protecting water quality as long as N fertilizer application is not too high, and provided that the grass component in mixture is sufficiently large (> 30%) so that nitrate absorption can occur before the water leaches out of the root zone (Eriksen et al., 2004). Indeed, with high application of N fertilizers (> 400 kg N ha\(^{-1}\) year\(^{-1}\)) or high legume proportion in swards (> 70%), considerable N leaching can be observed (Loiseau et al., 2001; Ledgard et al., 2009; Nyfeler, 2009). Thus, there are at least three reasons why N losses from legume-based grassland systems should be lower than from fertilized grass systems: (i) nitrogen is fixed symbiotically within the legume nodules and thus is not freely available in the soil in a reactive form, (ii) symbiotic N\(_2\) fixation activity is down-regulated if the sink of N for plant growth is small, and (iii) in balanced grass-legume mixtures, the grass roots take up N derived from legumes and from mineralization of soil organic matter.

Sustainable intensification and food security are required for a wide range of agricultural systems (Godfray et al., 2010) but the focus of action might differ among intensity levels of agricultural production. In less productive systems, such as those at the arctic sites in our study, savings in N fertilizer are probably of less importance due to low levels of N fertilizer use; however, our data show that for a given amount of N fertilizer input, higher N output (N\(_{\text{tot}}\) or forage protein per unit area) can be expected with grass-legume mixtures than with pure grass alone (more output for the same input). In systems that use high levels of N fertilizer to achieve high production levels, in contrast, the same N output can be achieved by mixed swards with less input of N fertilizer (the same output from less input), thereby
reducing energy use (Kitani et al., 1999), nitrate losses (Jensen et al., 2012), and GHG emissions (Davidson, 2009; Schils et al., 2013; Schmeer et al., 2014).

Finally, the use of grass-legume mixtures fulfills recent demands for climate change mitigation (Smith et al., 2013). The savings in application of N fertilizers that can potentially be achieved through the benefit of symbiotic N\textsubscript{2} fixation in grassland mixtures makes their use an important strategy for reduction of GHG emissions from agriculture (Smith et al., 2008; Canfield et al., 2010). Therefore, the use of such mixtures should be actively supported as they also meet major targets of practical grassland management, such as robustness in gains of total biomass yield despite variation in legume proportions, and applicability across wide environmental gradients (Suter et al., 2012; Finn et al., 2013). Provided that cultural barriers can be overcome (e.g. lacking knowledge in establishment of mixtures and their management to improve legume persistence), grass-legume mixtures are also a mitigation measure that can be implemented in the near future (Smith & Gregory, 2013; Smith et al., 2013). This is an important feature to meet the urgent needs for reductions in GHG emissions from the agriculture, forestry and other land use sector, which are estimated to be around 25% of total anthropogenic GHG output (Bustamante et al., 2014).

In conclusion, this study demonstrates that the N output of forage harvest is maximized in mixtures over a wide range of production levels. In the face of high economic and environmental costs of industrial N fertilizers (Kitani et al., 1999; Gruber & Galloway, 2008; Canfield et al., 2010), the contribution of symbiotic N\textsubscript{2} fixation by legumes to grassland N supply appears to be a key strategy to maintain and increase current levels of production and protein self-sufficiency in a more sustainable way than achieved so far.
Acknowledgments

We thank many colleagues, too numerous to mention, for technical assistance within this project. We are grateful to two anonymous referees for their comments on an earlier version of this paper. Coordination of this project was supported by the EU Commission through COST Action 852 ‘Quality legume-based forage systems for contrasting environments’. A contribution to the research leading to these results has been conducted as part of the AnimalChange project which received funding from the European Union’s Seventh Framework Programme (FP7/2007-2013) under the grant agreement no. 266018.
References


Supporting information

Additional Supporting Information may be found in the online version of this article:

Site information and details on the experimental design

Table S1 Site information on the sixteen sites including experimental details

Table S2 Sowing proportions of the two grass and legume species in experimental stands

Supporting text and information on the methods and analyses

Appendix A

Supplementary data

Table S3 Predicted total nitrogen yield of monocultures and mixing effects of the equi-proportional mixture

Table S4 Site values of environmental variables, total nitrogen yield \((N_{\text{tot}})\) and N yield gain in mixture \((N_{\text{gainmix}})\)

Table S5 Correlation matrix of the five environmental variables recorded in the experiment

Table S6 Regression estimates and their standard errors from the linear mixed model

Table S7 Summary of regressions relating realized legume proportion, total nitrogen yield \((N_{\text{tot}})\), and the relative N yield gain in mixture \((N_{\text{gainmix}}/N_{\text{tot}})\) to site productivity and four climatic variables

Fig. S1 Fitted lines to total N yield of individual sites as affected by sown legume proportion
### Tables

#### Table 1
Identity and functional types of species selected for experimentation in different regions of Europe. The species reflected agronomic importance and the adaptation to pedo-climatic conditions and were selected on the advice of local experts.

<table>
<thead>
<tr>
<th>Species group</th>
<th># Sites</th>
<th>Grass species</th>
<th>Legume species</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid European (ME)</td>
<td>12</td>
<td><em>Lolium perenne</em> L.</td>
<td><em>Trifolium pratense</em> L. <em>Trifolium repens</em> L.</td>
</tr>
<tr>
<td>Northern European (NE)</td>
<td>3</td>
<td><em>Phleum pratense</em> L.</td>
<td><em>Trifolium pratense</em> L. <em>Trifolium repens</em> L.</td>
</tr>
<tr>
<td>Other (O)</td>
<td>1</td>
<td><em>Lolium perenne</em> L.</td>
<td><em>Trifolium pratense</em> L. <em>Trifolium repens</em> L.</td>
</tr>
</tbody>
</table>

‡ Fast establishing species: fast germination and fast establishment, thereby providing adequate cover of soil in the sowing year and high biomass yields in the first and second years after sowing. These species often lack persistency.

§ Temporally persistent species: slower in germination and growth rate during establishment but highly competitive in the long run, therefore increasing in cover and biomass yields over initial years and constituting the majority of yield from the third year onwards.
**Table 2** Summary of analysis of total N yield \((N_{\text{tot}})\) regressed on legume proportion (Legume) in the multisite grassland experiment across Europe. *F*-tests refer to the fixed effects of the linear mixed model. See Material and Methods for detailed information on the model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Df(_{\text{num}})</th>
<th>Df(_{\text{den}})</th>
<th><em>F</em>-value</th>
<th><em>P</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Legume linear (Legume)</td>
<td>1</td>
<td>14.9</td>
<td>19.4</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Legume quadratic (Legume(^2))</td>
<td>1</td>
<td>15.0</td>
<td>19.4</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Legume cubic (Legume(^3))</td>
<td>1</td>
<td>14.9</td>
<td>12.0</td>
<td>0.003</td>
</tr>
<tr>
<td>Year</td>
<td>2</td>
<td>12.8</td>
<td>9.4</td>
<td>0.003</td>
</tr>
<tr>
<td>DeltaBiomass(^\dagger)</td>
<td>1</td>
<td>749.4</td>
<td>1664.4</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Legume x Year</td>
<td>2</td>
<td>10.4</td>
<td>5.6</td>
<td>0.022</td>
</tr>
<tr>
<td>Legume x DeltaBiomass</td>
<td>1</td>
<td>730.0</td>
<td>4.7</td>
<td>0.030</td>
</tr>
<tr>
<td>DeltaBiomass x Year</td>
<td>2</td>
<td>715.1</td>
<td>7.8</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

\(^\dagger\) DeltaBiomass is an effect on \(N_{\text{tot}}\) that is uncorrelated to the other predictors (see Appendix A, Supporting Information, for details). Seed density was not significant and was omitted. For regression estimates and their standard errors see Table S6, Supporting Information.

\(Df_{\text{num}}\): degrees of freedom of term; \(Df_{\text{den}}\): degrees of freedom of error (which can be fractional in restricted maximum likelihood analysis)
Table 3  Total nitrogen yield \( (N_{\text{tot}}) \) in monocultures and grass-legume mixtures and N yield gain \( (N_{\text{gain}}) \) in stands containing legumes across all sites of the continental-scale grassland experiment. Values are the fixed estimates (±1 standard error) derived from the linear mixed model corresponding to Table 2 (eqn. S1, Appendix A, Supporting Information, and Table S6). N yield gain in legume pure stands \( (N_{\text{gainleg}}) \) and mixtures \( (N_{\text{gainmix}}) \) was calculated as the difference between \( N_{\text{tot}} \) of a stand and \( N_{\text{tot}} \) of the pure grass stand. Values of \( N_{\text{tot}} \) in mixture \( (N_{\text{totmix}}) \), \( N_{\text{gainmix}} \), and \( N_{\text{gainmix}}/N_{\text{totmix}} \) are predicted for a stand with one third proportion of legumes (L) and two thirds of grasses (G).

<table>
<thead>
<tr>
<th>Pure stands</th>
<th>Mixture ( 1/3:2/3 ) L:G</th>
<th>95% of maximum ( N_{\text{tot}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grass</td>
<td>Legume(^\ddagger)</td>
<td>L:G</td>
</tr>
<tr>
<td>Year 1</td>
<td>178 (±25.8)</td>
<td>299 (±37.8)</td>
</tr>
<tr>
<td>Year 2</td>
<td>173 (±26.2)</td>
<td>211 (±36.4)</td>
</tr>
<tr>
<td>Year 3</td>
<td>119 (±16.4)</td>
<td>162 (±23.1)</td>
</tr>
</tbody>
</table>

\( (N_{\text{totmix}}), N_{\text{gainmix}}, \text{and } N_{\text{gainmix}}/N_{\text{totmix}} \) are predicted for a stand with one third proportion of legumes (L) and two thirds of grasses (G).

b) \( N_{\text{gain}} \) (kg ha\(^{-1}\) year\(^{-1}\))

<table>
<thead>
<tr>
<th>Grass</th>
<th>Legume(^\ddagger)</th>
<th>L:G</th>
<th>( N_{\text{gainleg}} )</th>
<th>( N_{\text{gainmix}} )</th>
<th>( N_{\text{gainmix}}/N_{\text{totmix}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>-</td>
<td>121 (±29.7)</td>
<td>108 (±22.7)</td>
<td>0.38 (±0.090)§</td>
<td></td>
</tr>
<tr>
<td>Year 2</td>
<td>-</td>
<td>38 (±24.6)</td>
<td>80 (±18.7)</td>
<td>0.32 (±0.085)</td>
<td></td>
</tr>
<tr>
<td>Year 3</td>
<td>-</td>
<td>43 (±12.1)</td>
<td>76 (±16.2)</td>
<td>0.39 (±0.097)</td>
<td></td>
</tr>
</tbody>
</table>

\( \ddagger \) Values for pure legume stands (sown as pure legume seeds) are at maximal realized legume proportion across sites, which were 100%, 97%, and 80% in years 1, 2, and 3, respectively (compare Fig. 4).

Figure legends

**Fig. 1.** Sites of the coordinated field experiment to analyze total N yield in grass-legume mixtures. Site numbers refer to Kirwan *et al.* (2014) and are as follows: 1: Belgium, 10: Germany_a, 11: Germany_b, 13: Iceland_a, 14: Iceland_b, 15: Ireland, 18: Lithuania_a, 20: Lithuania_c, 22: Norway_a, 23: Norway_b, 24: Norway_c, 26: Poland_a, 27: Poland_b, 34: Switzerland, 35: Wales_a, 36: Wales_b (see Table S1, Supporting Information, for further information on sites). Symbols refer to the set of plant species sown (see Table 1).

**Fig. 2.** Illustrative example of potential fractions of total nitrogen (N) yield in grass-legume mixtures and the respective monocultures. In grass monocultures (legume proportion = 0), the total N yield accumulated by the sward derives solely from the soil and from fertilizer N (quantity A). In legume monocultures (legume proportion = 1), the total N yield is greater (A + B), as legumes also have access to atmospheric N through symbiotic N$_2$ fixation. Mixing grasses with legumes should lead at first to a linear increase of accumulated N by the stand with increasing legume proportion (continuous white line) through a constant amount of N added to the system per unit of legumes. In addition, positive grass-legume interactions (e.g. stimulation of symbiotic N$_2$ fixation, N transfer) can result in a nonlinear surplus of N yield in mixtures (quantity C). The area of B+C defines the N yield gain (N$_{gain}$, see text) for varying legume proportions in mixtures as compared to grass monocultures.

**Fig. 3.** Total nitrogen yield (N$_{tot}$) of monocultures (Grass, Legume) and mixtures at three years for sixteen sites across Europe. N$_{tot}$ is averaged across seed density and additionally across monocultures of the two grass and legume species, respectively. Sites are arranged in order of decreasing means of monoculture N$_{tot}$ averaged over the first two experimental years, with horizontal bars denoting the annual mean N$_{tot}$ of mixtures. Inference on differences
between mixtures and *grass* monocultures is indicated on top of each panel (Wilcoxon rank sum test on the unpoled data). Missing symbols in b) for site 15 and in c) for sites 15 and 20 reflect the absence of legumes. Site numbers follow the codes used in Fig. 1, and missing site-years indicate no analysis of N concentration.

*** $P \leq 0.001$, ** $P \leq 0.01$, * $P \leq 0.05$, ns = not significant

**Fig. 4** Total nitrogen yield ($N_{\text{tot}}$) as affected by legume proportion in swards at three years of the multisite grassland experiment. The bold black line displays the predicted $N_{\text{tot}}$ across all sites (fixed mean from linear mixed regression corresponding to Table 2), and coloured lines display predicted $N_{\text{tot}}$ for individual sites (the variation around the fixed mean) over the range of legume proportion realized in the respective year. Numbers on lines refer to sites and are located at the mean legume proportion realized across all stands per site. The shaded area indicates N yield gain in mixture as compared to grass monoculture across sites (compare Fig. 2). The horizontal bold line at the bottom of the graph indicates the range of legume proportion for which $N_{\text{tot}}$ across all sites was significantly smaller ($P \leq 0.05$) than at maximum (■); consequently, the dotted line displays the non-significant range. No line could be drawn for site 15 in year two and for sites 15 and 20 in year three because legumes were absent. Other missing site-years indicate no analysis of N concentration.

**Fig. 5.** Correlation between legume proportion in the sward and the climatic variables minimum site temperature (a) and annual precipitation (b), and correlation between the relative N yield gain in mixture ($N_{\text{gain mix}}/N_{\text{tot mix}}$) and the same climatic variables (c, d), realized legume proportion (e), and site productivity (f). Data are site means across the first two experimental years, with $N_{\text{gain mix}}$ and $N_{\text{tot mix}}$ being estimated for a mixture with one third proportion of legumes and two thirds of grasses (following eqn. S2, Appendix A, Supporting Information). Pearson’s $R$, its significance, and the trend line following least square
regression refer to filled symbols (see Table S7, Supporting Information, for regressions on all climatic variables). Site 15 (Ireland, open symbol) was omitted from these analyses because legumes were absent from the second year onward. For site names to labels see Fig. 1.
(a) Year 1

(b) Year 2
Figure 1. Relationships between realized legume proportion and (a) minimum temperature, (b) precipitation, (c) realized legume proportion and productivity, (d) minimum temperature and productivity, (e) realized legume proportion and productivity, (f) productivity and productivity. The correlation coefficients (R) and p-values (P) are indicated for each relationship. The data points are labeled with their respective site numbers.