Design and assessment of legume-supported cropping systems

Moritz Reckling
Faculty of Natural Resources and Agricultural Sciences
Department of Crop Production Ecology
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2018

Acta Universitatis agriculturae Sueciae
2018:69

Cover: Co-learning with farmers, advisors and scientists on how to improve grain legume cropping systems (photo: G. Stange, ZALF)

ISSN 1652-6880
ISBN (print version) 978-91-7760-280-4
ISBN (electronic version) 978-91-7760-281-1

Print: SLU Service/Repro, Uppsala 2018

Design and assessment of legume-supported cropping systems

Abstract

Legumes provide high quality protein for food and feed as well as other ecosystem services, but it is still challenging to use them to meet the growing global demand for protein, partly because European farmers consider their cultivation unprofitable and risky. This thesis aims to *design* legume-supported cropping systems and *assess* their environmental and economic impacts along with their production risks in European agriculture.

The approaches used included (i) the development of a framework to design cropping systems and to assess impacts of management, (ii) modelling the impact of integrating legumes into cropping systems and assess trade-offs, (iii) the development of a statistical method to quantify crop yield stability independent of the mean yield, (iv) assessing grain legume yield stability statistically compared to other crops using data from long-term experiments, and (v) participatory methods to re-design legume-supported cropping systems.

The framework consists of a rule-based rotation generator and algorithms to calculate impact indicators, following a three-step approach: (i) generate rotations, (ii) evaluate crop production, and (iii) assess cropping systems. It was used to design and assess legume-supported cropping systems in five case study regions in Europe and to identify trade-offs between economic and environmental impacts. On average, the generated cropping systems with legumes reduced N_2O emissions by 18 % and 33 % and N fertilizer use by 24 % and 38 % in arable and forage systems, respectively, compared to systems without legumes. Grain legumes increased gross margins in two of five regions and forage legumes in all three study regions. A scale-adjusted coefficient of variation was developed as a stability measure that accounts for mean yield differences. Using data from five long-term experiments in northern Europe, this method showed that yield instability of grain legumes (30 %) was higher ($P < 0.001$) than that of autumn-sown cereals (19 %), but lower ($P < 0.001$) than that of other spring-sown broad-leaved crops (35 %), and only slightly greater ($P = 0.042$) than spring-sown cereals (27 %). The combination of on-station and on-farm trials with crop rotation modelling was useful when re-designing cropping systems. Nine agronomic practices were identified for improving grain legume production at the farm level.

In this thesis, it is shown that legumes can provide both economic and environmental benefits, the instability of yields is similar to other spring crops and that cropping systems can be re-designed effectively in a co-learning process with farmers.

Keywords: Diversification, participation, pulses, rotations, trade-offs, yield variability

Author's address: Moritz Reckling, SLU, Department of Crop Production Ecology, P.O. Box 7043, SE-750 07 Uppsala, Sweden.

E-mail: moritz.reckling@slu.se

Dedication

To my wife & children, and to my parents.

Contents

List of publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text:

I Reckling M.*, Hecker J.-M., Bergkvist G., Watson C., Zander P., Stoddard F., Eory V., Topp K., Maire J., Bachinger J. (2016). A cropping system assessment framework - evaluating effects of introducing legumes into crop rotations. *European Journal of Agronomy* 76:186-197.

II Reckling M.*, Bergkvist G., Watson C. A., Stoddard F.L., Zander P. M., Walker R., Pristeri A., Toncea I., Bachinger J. (2016). Trade-offs between economic and environmental impacts of introducing legumes into cropping systems. *Frontiers in Plant Science* 7:669.

III Döring, T.F. and Reckling, M.* (2018). Detecting global trends of cereal yield stability by adjusting the coefficient of variation. *European Journal of Agronomy,* 99:30-36.

IV Reckling, M.*, Döring, T. F., Bergkvist, G., Stoddard, F. L., Watson, C. A., Seddig, S., Chmielewski, F.-M., and Bachinger, J. (2018). Grain legume yields are as stable as other spring crops in long-term experiments across northern Europe. *Agronomy for Sustainable Development,* 38:63

V Reckling, M.*, Bergkvist, G., Watson, C. A., Stoddard, F. L. and Bachinger, J. Re-designing grain legume cropping systems using systems' agronomy. (submitted manuscript)

* Corresponding author.
Papers I-IV are published Open Access.

The contribution of Moritz Reckling to the papers included in this thesis was as follows:

I Designed the study and built the model with co-authors from ZALF and SRUC, analysed the data, and wrote the paper in collaboration with the co-authors.

II Designed the study, performed the modelling with data provided by co-authors, led the validation of the results and interpretation and wrote the paper in collaboration with the co-authors.

III Interpreted the results, prepared the figures and wrote the manuscript together with the first author, who performed the statistical analysis.

IV Designed the study, analysed the data provided by co-authors, and wrote the paper in collaboration with the co-authors.

V Designed the study, performed the analyses, and wrote the paper in collaboration with the co-authors.

List of tables

List of figures

Abbreviations

aCV	Adjusted coefficient of variation
BNF	Biological nitrogen fixation
C:N ratio	Carbon to nitrogen ratio
CO_2	Carbon dioxide
CS	Cropping system
CV	Coefficient of variation
DEED	Describe, explain, explore, design cycle
DF	Degrees of freedom
EF	Emission factor
EIP-AGRI	European Innovation Partnership in agriculture
EU	European Union
GxExM	Genotype, environment, management interaction
LTE	Long-term experiment
MASC	Multi-attribute assessment of the sustainability of CS
N	Nitrogen
N_2O	Nitrous oxide
NL lupin	Narrow-leafed lupin (*Lupinus angustifolius* L.)
PGRO	Processors and Growers Research Organisation
POLAR	Power law residuals
PRACT	Prototyping Rotation and Association with Cover crop and no Till
ROTOR	ROTations in Organic farming systems
SE	Standard error
TPL	Taylors Power Law

1 Introduction

Crop production in Europe is intensive, specialised and responsible for negative environmental impacts (Clark & Tilman, 2017). The combined effect of the intensification of livestock farming with political support for cereal production and imported soybean (Voisin *et al.*, 2002) led to a 71 % deficit in high-protein crop commodities in Europe in 2016 (Murphy-Bokern *et al.*, 2017).

While grain legumes were grown on 14.5 % of the global arable cropped area in 2014, they were grown on only 1.5 % in Europe (Watson *et al.*, 2017). Relative to cereals, grain legumes have relatively low yield levels and low yield stability as estimated from national yield data (Cernay *et al.*, 2015). A low market value compared to imported soybean products make growing of grain legumes less profitable than cereals in current supply chains (Preissel *et al.*, 2017; Meynard *et al.*, 2018). Farmers, grower organizations and experts from the European Innovation Partnership program (Von Richthofen *et al.*, 2006a; EIP-AGRI, 2014; Zimmer *et al.*, 2016; PGRO, 2018) hold low temporal yield stability responsible for the low proportion of grain legumes among other factors. Other drivers for the low proportion are associated with the specialization and intensification of farms on cereal crops, rapeseed and maize and importing relatively cheap protein for livestock feed (Zander *et al.*, 2016), unpredictable policy support for protein crops (Bues et al., 2013), and lack of awareness of the positive rotational effects of legumes at the cropping system (CS) scale (Preissel *et al.*, 2015). At the same time there is a high consumer willingness to pay for animal products produced with local protein feed (Profeta & Hamm, 2019), but this has not yet generated a large market for European grain legumes.

In Europe, perennial forage legumes as sole crop or in mixtures are grown on larger areas than grain legumes (Bues *et al.*, 2013). According to Phelan *et al.* (2015), the main advantages of forage legumes compared to other crops are the low reliance on N fertilizer and the high feed value. Disadvantages are the lower persistence than grass under grazing, high risk of livestock bloat and difficulty to conserve as silage or hay, and disease (Phelan *et al.*, 2015). The specialization of farming and associated spatial decoupling of livestock and crop production is one of the major constraint to forage legumes in Europe (Lemaire *et al.*, 2015) and limits their use in CS.

Forage and grain legumes represent one of the highest quality foraging resources for pollinators (Decourtye *et al.*, 2010). Their mass-flowering contributes to the maintenance of populations of wild and domesticated bees by providing nectar and pollen (Westphal *et al.*, 2003). While forage legumes have been found to increase population size and diversity of earthworms and Collembola (Eisenhauer *et al.*, 2009; Sabais *et al.*, 2011) there is less information on the effects of grain legumes on them. Grass-clover mixtures provide habitat for farmland birds such as skylark, corn bunting, yellow wagtail and whinchat especially when modified harvesting measures are implemented (Stein-Bachinger & Fuchs, 2012).

The integration of grain and forage legumes into agricultural systems has been identified as a way to increase its sustainability (Jensen *et al.*, 2012; Magrini *et al.*, 2016; Stagnari *et al.*, 2017; Watson *et al.*, 2017) and the increased consumption of legume-rich diets will provide health benefits for humans and livestock (Foyer *et al.*, 2016; Röös *et al.*, 2018).

Legumes have an advantage over other crops in that they can form symbiotic associations with nitrogen-fixing bacteria, making them self-sufficient in terms of nitrogen (N) acquisition (Peoples *et al.*, 2009a). Estimates of biological nitrogen fixation (BNF) vary widely and depend on the legume and rhizobia genotype, crop management, and environmental condition (Giller *et al.*, 2014). Under European conditions, BNF in faba bean (*Vicia faba* L.) can range between 73-335 kg N ha^{-1} $year^{-1}$ (Jensen *et al.*, 2010) and in white clover (*Trifolium repens* L.) between 10-545 kg N ha^{-1} $year^{-1}$ (Carlsson & Huss-Danell, 2003). Grain and forage legumes provide rotational services that influence the yield and quality of the subsequent crops. These agronomic benefits encompass N and non-N related preceding crop effects (Chalk, 1998) that are difficult to separate empirically. While the 'nitrogen effect' comprises the provision of N to the subsequent crops (Peoples *et al.*, 2017), the other benefits include the 'break-crop effect' that occurs when a disease cycle is broken (Robson *et al.*, 2002), benefits to soil organic matter and structure (Hernanz *et al.*, 2009) and phosphorus mobilisation (Shen *et al.*, 2011). These effects occur especially in cereal-dominated CS that are common in large parts of Europe. According to different meta-analyses of field experiments, cereal yields were 1.46 t ha^{-1} in temperate Europe (Preissel *et al.*, 2015), 1.2 t ha^{-1} in Australia, Europe and North America (Angus *et al.*, 2015) and 0.49 t ha^{-1} in

Sub-Saharan Africa (Franke *et al.*, 2018) higher after grain legumes than after cereal pre-crops. Due to the soil mineral N benefits derived from legumes to the subsequent crop (Peoples *et al.*, 2017), N fertilisation can be reduced by 23-31 kg ha^{-1} without reducing the yield benefit (Preissel *et al.*, 2015). Pesticides and soil tillage can also be reduced in crops following grain legumes because of the reduction in pathogen populations and improved soil structure (Von Richthofen *et al.*, 2006b; Angus *et al.*, 2015). Despite providing many benefits for the environment an cropping system, grain legume production is constrained by a large number of pests, diseases and weeds that limit the cultivation in the rotation, especially for pea (*Pisum sativum* L.) (Döring, 2015).

Cereal yields are also greater when grown in rotations with forage legumes (Persson *et al.*, 2008; Bergkvist & Båth, 2015; St-Martin *et al.*, 2017). While the effects are greater in year one and two after incorporation of the legume crop, Bergkvist and Båth (2015) found that oat yields were 0.3 t ha^{-1} greater even in the third season after the grass-clover compared to oat yields in a rotation without leys. The higher yields in rotations with forage legumes are considered to be related to an increase in soil organic matter (Johnston *et al.*, 2009), improved soil structure (Lüscher *et al.*, 2014) and larger amounts of soil mineral N (Müller-Stöver *et al.*, 2012). As perennial forages, they reduce the seed bank and storage organs of perennial weeds, thus reducing the need for other weed control measures (Håkansson, 2003).

While the general benefits and limitations of grain and forage legume cultivation in Europe have been reviewed extensively (Jensen *et al.*, 2012; Voisin *et al.*, 2014; Phelan *et al.*, 2015; Suter *et al.*, 2015; Zander *et al.*, 2016; Stagnari *et al.*, 2017; Watson *et al.*, 2017), these reviews have mostly focused on single crops and an annual perspective. An integrated assessment of legume-supported CS including agronomic, environmental and economic factors is needed to *design* novel systems and *assess* their impacts. Such a systems perspective will allow assessing the opportunities and challenges of legumes relative to each other and identifying avenues for an intensification of legume production in European agriculture.

1.1 Environmental impacts

Growing legumes contribute to climate change mitigation by reducing CO_2 emissions arising from N fertilizer production and nitrous oxide (N_2O) emissions from CS (Jensen *et al.*, 2012; Jeuffroy *et al.*, 2013). The effect of N_2O as a greenhouse gas is estimated at 265 times that of CO_2 on a 100-year basis and the intensification of agriculture has led to increased N_2O emissions

(Canfield *et al.*, 2010). The emissions from grain and forage legumes are generally lower than those from N-fertilized crops and pastures (Rochette & Janzen, 2005; Dusenbury *et al.*, 2008), because N_2O is derived from the nitrification of ammonia released from fertilizer or from mineralization of organic compounds and there are no emissions linked to the symbiotic N fixation process (Jeuffroy *et al.*, 2013). Nitrification produces nitrate, which is a prerequisite for denitrification that is converted to nitrite by nitrate reductases and the nitrite formed can undergo a variety of reactions. According to Canfield *et al.* (2010), the conversion to nitric oxide (NO) by nitrite reductases is the best known and probably occurs most frequently in arable soils. Losses occur both during the production of the crop and afterwards, the amount of the latter depending on residue management and whether the soil is left bare or a cover crop is sown (Watson *et al.*, 2017). From grain and forage legumes, Jensen *et al.* (2012) estimated N_2O-N emissions to average 1.3 kg ha^{-1} (ranging from 0.03-7.1 kg ha^{-1}), while the emissions from cereals, maize, canola and pasture were 3.2 kg ha^{-1} (ranging from 0.1-12.7 kg ha^{-1}). Jeuffroy *et al.* (2013) also found N_2O fluxes to be significantly lower for pea and unfertilized wheat (*Triticum aestivum* L.) than for fertilized wheat and fertilized oilseed rape. There is a risk of N_2O emissions from legume crop residues after the harvest due to their narrow C:N ratio. The N from the residues are available for rapid conversion to N_2O as well as for leaching but both Jeuffroy *et al.* (2013) and Pappa *et al.* (2011) found similar N_2O emissions compared to non-legume crops. Emissions are generally larger from soils with high N status and with larger applications of N fertilizer and manure (Rees *et al.*, 2012) because of greater nitrification of ammonia. Soil type, environmental factors and the interaction with crop management also affect emissions (Héault *et al.*, 2012; Graf *et al.*, 2016) because they determine the population composition of denitrifying microbes, the amount and speed of mineralization of nitrate and the N demand of the main and cover crops. Since emissions are subject to the management of the whole cropping system and can be expected during and after crop growth, it is important to measure and estimate emissions over a whole rotation instead of just one year (Dusenbury *et al.*, 2008).

There is an increased risk of nitrate leaching after grain and especially after forage legumes compared to after cereals due to the N-rich nature of legume residues. This is a particular problem on sandy soils and when the precipitation is high because this leads to higher amounts of drainage water which is a major determining factor of leaching (Beaudoin *et al.*, 2005; Benoit *et al.*, 2014). In a five-year rotational experiment in organic farming, Eriksen *et al.* (2015) found the highest nitrate-N leaching with maize and narrow-leafed lupin (*Lupinus angustifolius* L.) (NL lupin) (> 50 kg ha^{-1}), medium leaching with 2-4-year old grasslands (25-50 kg ha^{-1}) and low leaching when barley was grown with an undersown cover crop (<10 kg ha^{-1}). Cover crops can significantly reduce the amount of leaching as has been demonstrated in several studies (Beaudoin *et al.*, 2005; Plaza-Bonilla *et al.*, 2015). Under intensive agriculture in northern France, cover crops reduced leaching by 50 % at the annual and 23 % at the rotation scale (Beaudoin *et al.*, 2005). When designing systems with grain legumes, it is important to consider both the effect of the individual species and also the overall rotational system. Beaudoin *et al.* (2005) and Plaza-Bonilla *et al.* (2015) found a strong positive correlation between the proportion of grain legumes in rotations and the concentration of nitrate in drainage and soil water. Growing cover crops after grain legumes reduced the average nitrate-N leaching over a rotation with one grain legume from 32 kg ha^{-1} to 21 kg ha^{-1} and a rotation with two grain legumes from 52 kg ha^{-1} to 18 kg ha^{-1} (Plaza-Bonilla *et al.*, 2015).

There is good evidence on the impacts of legumes on the environment such as on N_2O emissions and nitrate leaching, but these have not been brought into a systems perspective with economic impacts at the cropping system level.

1.2 Design and assessment of cropping systems

Interactions between crops and the subsequent crops in a rotation affect the agro-economic and environmental performance. Nitrogen mineralization, nitrate leaching, greenhouse-gas emissions, pests, diseases and weeds, and eventual crop yield are all affected by crop management and the crop sequence (Dogliotti *et al.*, 2003; Bachinger & Zander, 2007; Detlefsen & Jensen, 2007). These interactions are particularly important when legumes are included because of their diverse impacts on all crops in the sequence (Peoples *et al.*, 2009b; Köpke & Nemecek, 2010; Jensen *et al.*, 2012). The production of legumes often generates lower gross margins than the production of cereals or oilseeds, but their rotational effects increase the gross margins of subsequent crops, which can have a compensatory effect on the system gross margin (Preissel *et al.*, 2017). When Preissel *et al.* (2017) compared rotations with and without legumes in a European wide assessment, they found that 30 out of 45 rotations including legumes were economically competitive with rotations without legumes.

To consider the rotational effects of legumes in an environmental and economic assessment, the cropping system should be the basis for the analysis. The cropping system is defined in this thesis as a fixed, cyclical sequence of crops (rotation) with a determined length (Castellazzi *et al.*, 2008) and the management activities for each crop including soil tillage and fertilizer inputs, and the production orientation (arable or mixed).

There are different concepts for the design and evaluation of CS. Based on a literature review, Eckersten (2017) identified the factors to be defined in CS assessments, the systems to include (such as plant, land, management), the choice of the evaluated outputs (such as yield, ecosystem services) and the factors influencing these outputs (such as rotational effects, climatic conditions, prices). Bergez *et al.* (2010) proposed an operational four-step approach to design and assess CS: (i) generation, (ii) simulation, (iii) evaluation, and (iv) comparison and choice. The *generation* or formulation of crop rotations can follow either a quantitative or qualitative approach or a combination of both. Quantitative approaches generally use statistical data and many studies on crop rotations in Europe use data from the integrated administration and control system of the EU and its land parcel information system (Schönhart *et al.*, 2011; Steinmann & Dobers, 2013). The use of statistical data is often combined with mathematical approaches (Castellazzi *et al.*, 2008). Qualitative approaches generally focus on historical crop rotation systematizations or expert-derived rotations (Stein & Steinmann, 2018). The combination of quantitative and qualitative information to design crop rotations is the most common method in recent years. Studies combine statistical data with qualitative rules to derive a crop sequence typology (Schönhart *et al.*, 2011; Lorenz *et al.*, 2013; Stein & Steinmann, 2018), or they use rule-based models for specific conditions such as PRACT for conservation agriculture (Naudin *et al.*, 2015) or ROTOR for organic farming (Bachinger & Zander, 2007). While statistical data represent current farming trends that are influenced by current policy and market drivers, they do not allow investigation of the effects of novel crops or management strategies. Rule-based models have been developed to generate novel systems based on a combination of quantitative and qualitative criteria.

The *evaluation* of CS can be based on *simulation* with dynamic models or assessment with static models. With dynamic crop models it is possible to simulate soil-crop processes in detail, but their high demand on data limits their application (Jones *et al.*, 2017). They have other limitations, such as an inability to generate systems or simulate the processes of crop rotations, and few are calibrated for grain legumes (Kollas *et al.*, 2015). Static models allow the assessment of rotational effects (Bachinger & Zander, 2007; Naudin *et al.*, 2015) without the necessity to disentangle the biotic and abiotic processes involved. They require less input data than dynamic models, and can combine crop rotation generation and evaluation. Static models provide less detailed output than dynamic models, and for evaluating rotations they need input from

qualified experts on crop management, rotational effects and crop rotation design (Naudin *et al.*, 2015).
For the *comparison and choice* step, following Bergez *et al.* (2010), multi-criteria analysis can be used that takes conflicting objectives underlying the economic, social and environmental dimensions of sustainability into account (Sadok *et al.*, 2009b). It is used to identify potential trade-offs between impacts using multi-criteria decision-aid methods (Sadok *et al.*, 2009a; Sadok *et al.*, 2009b; Carof *et al.*, 2013) that facilitate the discussion with stakeholders.
Overall, the design and evaluation of CS benefits from applying models. However, to support participatory problem-solving and co-design processes, these models must be both sophisticated enough to be relevant and sufficiently simple and flexible to be applied in very different settings and for different questions (Ditzler *et al.*, 2018). Static models meet these requirements, but so far there is not one model that evaluates legumes in CS for a wide range of environmental and farming conditions, and assessing economic and environmental impacts in a multi-criteria analysis.

1.3 Assessment of temporal yield stability

In addition to the rotational perspective in the CS analysis, there is a temporal dimension of the productivity and sustainability of legume crops. In agricultural sciences, the concept of stability is used as a criterion to measure the spatial or temporal constancy of specific features of agricultural systems (Urruty *et al.*, 2016). In the face of climate change and to achieve global food security, the stability of agricultural systems is becoming as important as their productivity (Olesen *et al.*, 2011; Kalkuhl *et al.*, 2016; Knapp & van der Heijden, 2018; Najafi *et al.*, 2018). Temporal (interannual) yield stability is influenced by the crop genotype (G), the management (M) of the CS and the biophysical environment (E), and their interactions as expressed by the relation GxExM (adapted from Giller *et al.,* 2014).
There are two main contrasting concepts of yield stability, namely the static and the dynamic (Becker & Léon, 1988). In the static or variance-based concept, the most stable genotype maintains a constant yield across environments while the dynamic or regression-based concept implies for a stable genotype a yield response in each environment that is always parallel to the mean response of all tested genotypes (Annicchiarico, 2002). While stability analysis was originally used to assess the stability of crop genotypes across environments, the analysis of yield stability of CS (Piepho, 1998) and changes over time due to climate change has gained importance (Lobell *et al.*,

2011b; St-Martin *et al.*, 2017; Hoffmann *et al.*, 2018; Tigchelaar *et al.*, 2018; Webber *et al.*, 2018).

Global average temperatures have risen by roughly 0.13°C per decade since 1950 and temperature trends from 1980 to 2008 exceeded one standard deviation of historic year-to-year variability (Lobell *et al.*, 2011b). The increased climate variability is also associated with a decreased stability of crop yields (Lobell & Field, 2007; Peltonen-Sainio *et al.*, 2010; Trnka *et al.*, 2012; Ray *et al.*, 2015; Najafi *et al.*, 2018; Tigchelaar *et al.*, 2018).

There are only very few studies that have quantified yield stability of grain legumes (Hawtin & Hebblethwaite, 1983; Peltonen-Sainio & Niemi, 2012; Cernay *et al.*, 2015). In agricultural statistics, yields of grain legumes were found to be less stable than other crops in Europe (Cernay *et al.*, 2015). Existing studies provide no explanation about the causes for the low stability. Therefore, the assessment of yield stability is particularly relevant in European-grown grain legumes. The suggested causes for low yield stability in grain legumes can be associated with, (i) the indeterminate growth habit that allows the crop to respond to good conditions such as high water availability and adequate temperature (Pilbeam *et al.*, 1990) or to stop growing and reproducing in poor conditions (Stoddard *et al.*, 2006), (ii) BNF that affects yield and can be reduced or fail in poor conditions (Kermah *et al.*, 2018), and (iii) low investment in breeding for yield, disease resistance and stress tolerance (Magrini *et al.*, 2016), which could influence the hardiness of plants when confronted with stresses and lead to lower yield stability.

To assess yield stability, many different regression- and variance-based indicators have been proposed (Eberhart & Russell, 1966; Becker, 1981; Becker & Léon, 1988; Huehn, 1990; Eghball & Power, 1995; Piepho, 1998; Dehghani *et al.*, 2008). There is also an extensive literature that deals with the comparison of various stability indicators (e.g. Becker & Léon, 1988; Crossa, 1988; Ferreira *et al.*, 2006). All of these indicators are less suitable for the comparison of yield stability between grain legumes and other crops because of the large differences in mean yield that are not reflected in the analysis. In addition, interpretation of the results of the stability analysis is difficult because different indicators may lead to contrasting conclusions reflecting different concepts of stability (Dehghani *et al.*, 2008) and the complex calculations involved makes it difficult to separate ‘true’ effects from mathematical artefacts.

A relatively simple stability indicator is the coefficient of variation (CV) that is one of the most frequently used in agronomic and ecological research (Francis & Kannenberg, 1978; Küchenmeister *et al.*, 2012; Ray *et al.*, 2015; Di Matteo *et al.*, 2016; St-Martin *et al.*, 2017; Knapp & van der Heijden, 2018; Müller *et* *al.*, 2018). The CV is defined as the standard deviation σ divided by the mean μ, and is expressed as a percentage of the mean: CV = (σ/μ) · 100 %. The calculation of the CV takes into account when the standard deviation increases with the mean by dividing it by the mean. Applying the CV implies the assumption that the standard deviation increases linearly with the mean. However, Döring *et al.* (2015) showed that under certain conditions, the unguarded interpretation of the CV of crop yield data may be misleading, especially when the crop yield data spans a large numeric range. This is often the case when yield data from various CS, locations and crop species are compared or when yield data is analysed over long time periods that include an increase in mean yields. In these cases, the CV of crop yield data tends to decrease with increasing mean (Döring *et al.*, 2015; Knapp & van der Heijden, 2018). This is because the yield data frequently follows a specific power-law relationship between the sample variance σ^2 and the sample mean μ. This power-law relationship, $\sigma^2 = A\mu^b$, is known as Taylor's Power Law (TPL). Logarithmic transformation of TPL results in a linear relationship, expressed as the equation $\log(\sigma^2) = a + b \log(\mu)$ with $a = \log(A)$. TPL was described by the British ecologist Roy Taylor (Taylor, 1961) who mathematically derived the relationship and a method to transform such data. TPL has since then been applied in hundreds of data sets from population ecology (Cohen *et al.*, 2012; Cohen *et al.*, 2013) and other sciences (Eisler *et al.*, 2008). Since TPL has also been found to hold well in crop yield data (Döring *et al.*, 2015; Reckling *et al.*, 2015; Knapp & van der Heijden, 2018), caution is needed when interpreting the CV or any of the stability indicators of crop yields. In most crop yield data sets analysed, a TPL-like relationships between mean and variance has been found (Döring *et al.*, 2015) and the CV systematically decreases nonlinearly with increasing mean.

With any of the existing yield stability indicators, there remain two methodological challenges in the assessment of yield stability: The risk of dependence on mean yield and the risk of a scale- and aggregation bias.

Since grain legumes are generally grown on smaller areas than cereals and other crops, aggregated yield data from national statistics is less suitable for the comparison of grain legume yield stability with other crops. Yield data from long-term field experiments (LTE) might be more suitable because they offer yield data of legumes and other crops under relatively controlled conditions over long time periods (Stützel *et al.*, 2016).

To assess whether yields of grain legumes are more or less stable than those of other crop species, a yield stability indicator needs to be developed by adjusting the standard CV to remove the dependence from the mean yield, and field-level yield measurements are required to avoid aggregation biases.

1.4 Re-design of cropping systems with farmers

The production of legume-based protein in Europe can be increased by increasing yields of grain legumes and by increasing the area used for their production. This can be achieved by the development of value chains to create markets for more protein products (Meynard *et al.*, 2018) and by improving the agronomy of grain legumes. The main challenge here is an agronomic advancement with practical impact and local adaptation by involving farmers. This can be achieved through co-design that is defined as involving different stakeholders in the collective exploration of solutions to a common problem and seeks to build and maintain a shared conception of the design problem to allow collaboration (Berthet *et al.*, 2018).

According to Doré *et al.* (2011), agronomic research in the past has focused mainly on outputs from simulation studies and statistical hypothesis testing with empirical data from experiments that were mainly conducted on experimental stations. The results clearly advanced agricultural research and knowledge on CS processes, but impacts were often far from the reality of farmers' fields (Doré *et al.*, 2011). Participatory methods have a long history in agricultural research (Farrington & Martin, 1988), but their use in the design of CS is continuously developing (Sinclair, 2017; Berthet *et al.*, 2018) and has not been sufficiently used for achieving practical impact. Doré *et al.* (2011) suggested that in order to advance CS research, more sources of knowledge need to be utilised, (i) making use of recent advances in plant sciences e.g. in crop modelling, (ii) learning lessons from the functioning of natural ecosystems e.g. from experiments and (iii) making more use of local farmers' knowledge through on-farm trials, surveys and participatory research.

Models are often used in purely academic design studies. Exceptions are FarmDESIGN (Groot *et al.*, 2012) that was used for the work with framers, producing results on productivity and efficiencies at the farm scale. Overall, the potential of using models in co-design has been insufficiently explored. Decision support tools can be used for the design and assessment of systems in various contexts e.g. PRACT (Naudin *et al.*, 2015) for designing rotations with cover crops and no till for conservation agriculture with smallholder farmers in Sub-Saharan Africa, MASC (Sadok *et al.*, 2009b) for the sustainability assessment with stakeholders, and ROTOR (Bachinger & Zander, 2007) for the design of organic CS with legumes together with farmers and advisors.

There are different options for participation of farmers in agronomic experiments (Catalogna *et al.*, 2018). Following the co-design approach, farmers are involved in formulating the research questions and hypotheses that are tested in on-station experiments, and in the interpretation of the results

during field days. Another option is when farmers experiment on their own farms (Catalogna *et al.*, 2018). Farmer-managed trials are a key element in co-design, by involving farmers, scientists and other stakeholders from the design stage, to the evaluation stage (Sumberg *et al.*, 2003; Sinclair, 2017; Catalogna *et al.*, 2018). Examples of on-farm trials are the experimentation with organic and agroecological practices (Catalogna *et al.*, 2018), the evaluation of crop cultivars (Schmidt *et al.*, 2018), the adaption of CS to climate change (Bloch *et al.*, 2015) and the unravelling of the causes of variability in crop yields and treatment responses with examples mainly from Africa (Falconnier *et al.*, 2016; Franke *et al.*, 2016; Ronner *et al.*, 2016; van Vugt *et al.*, 2018).
The DEED (Describe, Explain, Explore, Design) research cycle is a general conceptual framework for the design of CS which operationalizes systems agronomy (Giller *et al.*, 2015). This involves participatory work with farmers, modelling and experimentation. The DEED cycle supports the understanding of the complexity of farming and the generation of tailored options to re-design the CS of individual farmers. The cycle consists of four steps: (i) Describe current production systems and their constraints, (ii) Explain the consequences of current farm management, (iii) Explore options for agro-technological improvement and (iv) Design improved management systems (Giller *et al.*, 2015). The cycle is used for co-learning by farmers, advisors and agronomists, to identify which options fit best. Thus it provides a farm-specific solution rather than a 'silver bullet' (Giller *et al.*, 2011) by using a combination of methods available to agronomists such as crop rotation modelling, on-station and on-farm research (Descheemaeker *et al.*, 2016). The involvement of the actors in all steps of the cycle supports the local relevance of the designed options (Falconnier *et al.*, 2017; Sinclair, 2017). While such participatory work is common with smallholder farmers in Sub-Saharan Africa and Latin America (Dogliotti *et al.*, 2014; Descheemaeker *et al.*, 2016; Falconnier *et al.*, 2017; Ronner, 2018), it has been far less used in the development of agriculture in Europe (Prost *et al.*, 2018).
Co-design has the potential for improving the agronomy of grain legumes with farmers to achieve the needed agronomic advancement related to yield and other services and to result in practical impact and local adaptation.

2 Thesis aim and objectives

The aim of the thesis was to *design* legume-supported cropping systems and *assess* environmental and economic impacts along with production risks in European agriculture.

The overall aim was split into five specific objectives in separate papers:

1. Develop a cropping system assessment framework using a static and rule-based approach considering crop rotations and rotational effects (Paper I).

2. Assess the economic and environmental effects of integrating legumes into cropping systems, in order to identify potentials and limitations of increasing legume cultivation in Europe (Paper II).

3. Develop a novel method to quantify yield stability by adjusting the standard CV such that dependence from the mean yield is removed (Paper III).

4. Assess whether yields of grain legumes are more or less stable than those of other crops using field-level data from long-term field experiments from northern Europe and by accounting for Taylor's Power Law (Paper IV).

5. Describe and explain farmers' perceived constraints and opportunities for grain legume production, explore technical options at the field scale, re-design cropping systems and evaluate the role of different methods in agronomy in participatory research (Paper V).

3 Materials and methods

3.1 The work in the context of systems analysis

In this thesis the concept of systems analysis is used as a way to consider, in a balanced integration, the biophysical, economic, social and institutional aspects of the system under study (van Ittersum *et al.*, 2008). To facilitate the integrated analysis of agricultural systems a selection of methods from systems analysis was used, i.e., modelling, experiments, statistics and participatory approaches (Table 1). In Paper I, a CS model was developed to generate and evaluate CS in a participatory process with advisors and agronomists. The model was then used to identify trade-offs between economic and environmental services in five case study regions (Paper II). In Paper III, a scale-adjusted yield stability indicator was developed to account for differences in mean yield. It was applied to published cereal yield data to detect trends in changes of yield stability over time (Paper III) and to assess yield stability of grain legumes compared to other crop species using yield data from LTEs in Paper IV. Finally, a participatory research cycle was implemented to re-design CS with grain legumes using surveys, crop modelling, field experiments and simplified on-farm trials (Paper V).

Table 1. Selection of methods from farming systems analysis applied in this thesis

	Modelling	Experiments	Statistics	Participatory methods
Paper I	X			X
Paper II	X			X
Paper III			X	
Paper IV		X	X	
Paper V	X	X	X	X

The methodological approaches in the five papers cover different points in the design process from the assessment of CS to the design of new practices and span over different spatial scales or system boundaries from the field level in experiments to the farm scale in the participatory study. In Figure 1 they are organised along these two dimensions.

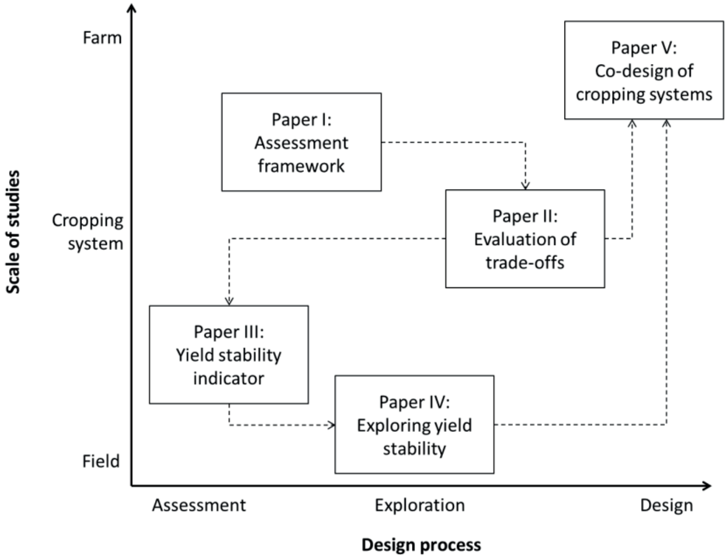

Figure 1. The allocation of Paper I-V in two dimensions, from assessment of cropping systems to design, and from field to farm scale. The dashed arrows show the connection between the papers along the two dimensions of the graph.

The developed framework (Paper I) was used to assess the environmental and economic impacts at the CS scale using data from the field scale and allowing interpretations at the farm scale. In paper II the framework was applied to assess CS and motivated the co-design at the farm level (Paper V), and the development and application of the yield stability indicator at the field level (Paper III-IV). In Paper V practices were identified that potentially increase the yield stability of grain legumes (quantified in Paper IV) and options for improving the economic and agronomic impacts were explored (identified in Paper II) by re-designing CS with a focus on farm-scale considerations.
Knowledge from the integrated assessment on different scales contributed to the exploration and co-design to achieve practical impact of results at the farm scale.

3.2 The cropping system framework (Paper I-II)

With the developed static and rule-based framework, it is possible to assess the impacts of CS on a set of environmental and economic indicators taking rotations and differences in crop management into account (Paper I). The framework follows three main steps: i) *generate crop rotations* using a rule-based rotation generator, ii) *calculate the impact of crop production activities* using static environmental and economic indicators, and iii) *assess and compare cropping systems*. The indicators were used in multi-criteria analyses for the identification of trade-offs between economic and environmental indicators in Paper II. The framework was applied to assess the impacts of introducing legumes into CS but other changes in CS can also be assessed.

3.2.1 Generate crop rotations

Crop rotations are generated with a rotation generator that produces fixed and cyclical 'agronomically sound rotations' for arable fields, following a fixed set of site-specific agronomic rules: (i) crop to crop suitability, (ii) maximum frequency of a crop in the rotation, (iii) minimum break between the same crop and (iv) maximum frequency of crops of the same crop type. Agronomists define the crops intended to be assessed such as crops currently grown and potential novel crops and then define the restriction values. The rotation generator combines all defined crops to produce all possible 3- to 6-year sequences following the restriction criteria and removes all possible duplicates and multiples. In an iterative process, the generated rotations are evaluated by experts, the restriction values modified e.g. when common rotations are not included, and a new set of rotations generated. Particular rotations like the *current farming* rotation ('business as usual') can be added manually to allow comparisons that are of particular interest.

3.2.2 Evaluate crop production

Single crop production activities (CPA) are assessed on an annual basis. Each CPA includes parameters of the main crop, the preceding crop and the site-specific management. CPAs are stored in a database with management data generated from structured surveys among 2-4 experienced agronomists per study region and include input data, crop yield, and management characteristics. Rotational effects are considered on a pre-crop basis. Experts estimate the expected influence of one crop on the next, i.e. the preceding crop effect in kg ha^{-1} of grain and forage yield, and the differences in fertilization and agro-chemical applications for different pre-crop types.

The following indicators are assessed for each CPA:

- Nitrate leaching is calculated as nitrate-N based on the soil type, preceding crop and crop management as a function of the soil leaching probability and the N surplus (see Paper I for details).
- Nitrogen fertilizer efficiency is calculated as the ratio of the N output in harvested grain or biomass to the N input from mineral and organic fertilizer.
- Nitrogen fertilizer use is calculated as the N applied from organic and mineral N fertilizer.
- Nitrous oxide emission from crop cultivation is calculated with the IPCC 2006 Tier 1 methodology (IPCC, 2006), including direct and indirect emission from fertilizer, manure and crop residues.
- Gross margins are calculated by subtracting variable costs from the revenues. Fixed and labor costs, subsidies, interest and cost of insurances are not taken into account.
- Infestation risks are assessed for each crop concerning selected pests, diseases and weeds that are problematic for production and influenced by sequence.

3.2.3 Assessment of cropping systems

In step 3, the generated crop rotations for each site (step 1) are combined with the evaluated CPAs (step 2) and impacts calculated on a per-hectare and per-year basis. The evaluated CS are then grouped depending on the research objective e.g. into ‘arable systems’ or ‘mixed systems’, ‘with legumes’ or ‘without legumes’. The outputs of the CS can then be analyzed e.g. by plotting environmental impacts against economic impacts or by comparing the mean impact between different systems.

3.2.4 Application of the framework to legumes

For this thesis, the framework was used to assess arable and mixed CS with and without legumes and was first tested in two case study regions (Paper I), namely Brandenburg in north-eastern Germany and Västra Götaland in south-western Sweden (Figure 2). It was then applied in three more regions across Europe (Paper II), eastern Scotland in the United Kingdom, Calabria in southern Italy, and Sud-Muntenia in Romania (Figure 2). The regions have contrasting climatic conditions and CS, and were selected to represent a broad range of bio-physical and socio-economic conditions and possible roles of legume production.

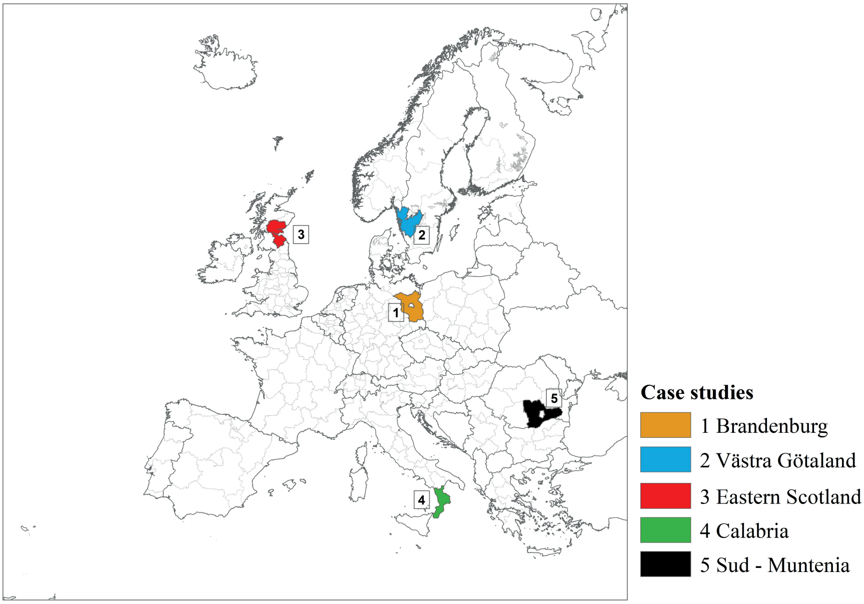

Figure 2. The case study regions at NUTS 2 level across Europe (Paper I-II). This map includes data produced by European National Mapping and Cadastral Agencies @ EuroGeographics.

Arable systems comprised only annual grain crops including grain legumes while mixed systems comprised grain and at least one forage crop i.e. temporary grass, silage maize and forage legumes in mixture or as sole crops. The data on crops and rotations were provided by 2-4 experts in each case study region.

Trade-offs between economic and environmental impacts were analysed using multi-criteria analysis based on the calculated indicators. To allow relative comparisons, impacts were normalized for each land capability and arable and mixed systems separately by dividing the impact of each single CS by the overall mean. The following CS were compared to evaluate trade-offs between economic and environmental impacts:

(1) *current farming* without legumes
(2) *economic-environmental optimized* systems without legumes
(3) *economic-environmental optimized* systems with legumes

Current farming without legumes represents the most common CS based on the crop proportions and expert knowledge. *Economic-environmental optimized* CS with and without legumes were selected from the large range of systems according to the gross margin (equal to, or up to 50€ lower than the system with the highest gross margin), nitrate leaching and N_2O emission (equal to or lower than the system with the highest gross margin).

3.3 Assessment of yield stability (Paper III-IV)

A scale-adjusted coefficient of variation (aCV) was used that removes the dependence of CV on the mean. The aCV was developed in Paper III and tested to assess changes in cereal yield stability over time and applied in Paper IV to quantify yield stability of grain legumes and other species.

3.3.1 The scale-adjusted coefficient of variation (aCV)

For the calculation of the aCV, a series of yield data over several years is used to estimate the mean ($\hat{\mu}$) and variance ($\hat{\sigma}^2$) over pre-defined periods. The *n* pairs of means $\hat{\mu}_i$ and variances $\hat{\sigma}^2{}_i$ (with index *i* from 1 to *n*) are subsequently used for the yield stability calculation.
For comparison with the adjusted coefficient of variation, the *standard* coefficient of variation CV_i is calculated as

$$CV_i = \frac{\hat{\sigma}_i}{\hat{\mu}_i} \cdot 100\,\% \qquad \text{(eqn. 1)}$$

The adjustment of the coefficient of variation follows four steps. First, following TPL, a linear regression is calculated for $\log_{10}$ of the variance over the $\log_{10}$ of the mean of all crops (Döring *et al.*, 2015) (example in Figure 3).

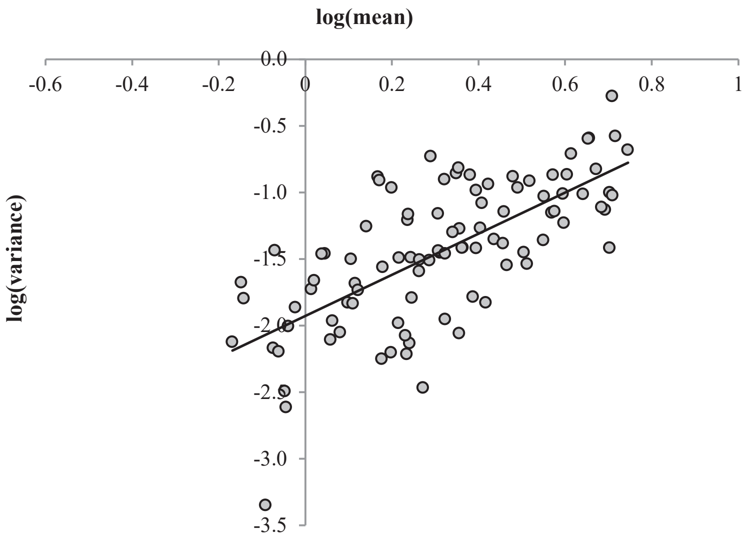

Figure 3. Linear regression of $\log_{10}$ of mean against $\log_{10}$ of variance for global yield data for rye, following Taylor's Power Law (Paper III). The regression follows y = a + bx with a = -1.93 ± 0.07 SE and b = 1.55 ± 0.17 SE (df= 93, Adjusted R^2 = 0.44, P < 0.001).

With $v_i = \log(\hat{\sigma}^2_i)$ and $m_i = \log(\hat{\mu}_i)$, the linear regression is $v = a + bm$. Second, the residuals u_i from this regression line, i.e. the POLAR (Power Law Residuals, Döring *et al.* 2015), are calculated as

$$u_i = v_i - (a + bm_i) \quad \text{(eqn. 2)}$$

Third, to account for the systematic relationship between the logarithm of the sample variance and the logarithm of the sample mean described by Taylor (1961), the logarithm of the variance is adjusted and subsequently used for calculating the coefficient of variation. The adjusted logarithm of the variance $\tilde{v}_i$ is

$$\tilde{v}_i = 2m_i + (\mathrm{b} - 2)\bar{m} + a + u_i \quad \text{(eqn. 3)}$$

where $\bar{m} = \frac{1}{n}\sum m_i$. The fourth and final step is to use the adjusted logarithm of the variance for calculating the adjusted coefficient of variation aCV_i.

$$aCV_i = \frac{\sqrt{g^{\tilde{v}_i}}}{\hat{\mu}_i} \cdot 100\,\% \quad \text{(eqn. 4)}$$

When the TPL regression slope b is < 2, the *standard* CV decreases non-linearly with increasing mean. In this case, $CV_i = \hat{\mu}_i^{\frac{b}{2}-1} g^{\frac{a}{2}} \cdot 100\,\%$, where g is the basis of the logarithm (Döring *et al.*, 2015). For adjusting the coefficient of variation, the dependence of the CV from the mean is removed, so that the new slope b_{adj} between log(mean) and adjusted log(variance) equals $b_{adj} = 2$ (see eqn. 3); as a consequence, $\hat{\mu}_i^{(b_{adj}/2-1)} = \hat{\mu}_i^{0} = 1$. This is visualized in Paper III.

3.3.2 Data sets

In Paper III, national yield data stored in the Food and Agriculture Organization database FAOStat (FAOSTAT, 2014) for wheat and rye (*Secale cereal* L.) were analysed to represent crop yield variation at a national level across the globe. Only countries with a complete set of yield data over the fifty year period (1964-2013) were used in the analysis. Further exclusion criteria were considered as explained in Paper III. For wheat and rye, these filters resulted in a final dataset of crop yields from 77 and 19 countries, respectively. The yield data were grouped into five consecutive decades (1964-1973, 1974-1983 etc.). Means ($\hat{\mu}$) and variances ($\hat{\sigma}^2$) were calculated per country within each decade and used for the analysis of yield stability.

In paper IV, data from five LTEs were used to quantify yield stability of grain legumes compared to other species. Long-term experiments from the United Kingdom, Sweden and Germany (Figure 4) were chosen and the yield data over 11-56 years was used to estimate the mean ($\hat{\mu}$) and variance ($\widehat{\sigma}^2$) over 8-year periods representing the crop rotation length. In total, the experiments provided 3768 site-year combinations. With regard to crop species and crop groups, data sets were balanced, i.e. at any one site, each crop species and crop group was grown every year, allowing analysis within sites and between sites.

Figure 4. The experimental sites across northern Europe (Paper IV). This map includes data produced by European National Mapping and Cadastral Agencies @ EuroGeographics.

Groups of crops were defined as 'grain legumes' including faba bean, pea and lupins (yellow and narrow-leafed), 'cereals', with spring barley, spring oat, spring wheat, winter barley, winter rye, winter spelt, winter triticale and winter wheat, and 'broad-leaved crops' with potato, sugar beet and winter oilseed rape. 'Spring-sown crops' included potato, sugar beet, grain legumes and spring cereals, and 'autumn-sown crops' winter cereals and winter oilseed rape.
Experiments include a diversity of CS and management systems to compare yield stability of different crop species. Sites differed in soil texture with clay content varying from 3 % to 25 %, annual precipitation from 545 mm to 667 mm and annual mean temperature from 8.3 to 10.1 °C.

3.3.3 Statistical analysis

The statistical analysis of data (Paper III-IV) was performed with the R software version 3.3.1 (R, 2016). In Paper III, the data were detrended following Ray *et al.* (2015) for each decade and each country, i.e. linear trends within each decade were removed so as to avoid penalising intended yield increases over time due to agronomic and technical progress, which would have led to higher variances. In Paper III and IV, the CV and aCV were tested for normal distribution. For testing the relationship between the $\log(\hat{\sigma}^2)$ and $\log(\hat{\mu})$, the CV and the mean yield, and the aCV and the mean yield, a linear model was applied with the *lm* function in R.

In paper III, the aCV was tested for significant changes over time by a linear mixed effects model with the *lme* function in R, using country as a random factor and decade as a fixed factor (Piepho *et al.*, 2003). In paper IV, significant differences between the groups of crops were tested with a linear mixed effects model with the *lme* function in R, using site as a random factor and crop group as a fixed factor.

3.4 Re-designing cropping systems (Paper V)

In Paper V, a participatory study was implemented with farmers, advisors and scientists that were part of or associated with two large grain legume demonstration networks supported through the German protein crop strategy, the soybean network www.sojafoerderring.de and the lupin network www.lupinen-netzwerk.de. The networks comprised a total of around 100 farms across Germany, of which 25 farms were involved in this study. The organic farms were medium- to large-scale and mainly mixed with arable and livestock activities. Farmers were interested and open to innovations, participated in regular monitoring of their fields and activities and explored alternative practices through testing technical options at field scale.

The study has a focus on NL lupin as an already well established crop in the region and soybean (*Glycine max* (L.) Merr.) as a potential novel crop.

3.4.1 Study area

Northern Germany was selected as the study area (Figure 5) because the analysis in Paper II revealed large trade-offs between economic and environmental impacts when growing grain legumes in Brandenburg, which is within the study area. The study area can be divided into an eastern part that is characterized by mostly sandy soils and low average annual rainfall of around 500 mm, and a western part with often better soils and larger average annual rainfall of around 700 mm. Grain legumes were cultivated on 126 500 ha in the study area in 2016, representing 0.8 % and 2.4 % of the total arable land in the western and eastern parts, respectively in 2017 (DESTATIS, 2017). NL lupin was grown on 27 100 ha and soybean on 2 800 ha in the entire study area in 2017 (DESTATIS, 2017). On the 25 farms, simplified on-farm trials were established and data collected for the crop rotation modelling and on farmers' opportunities and constraints, and field experiments were established on one experimental research station (Figure 5).

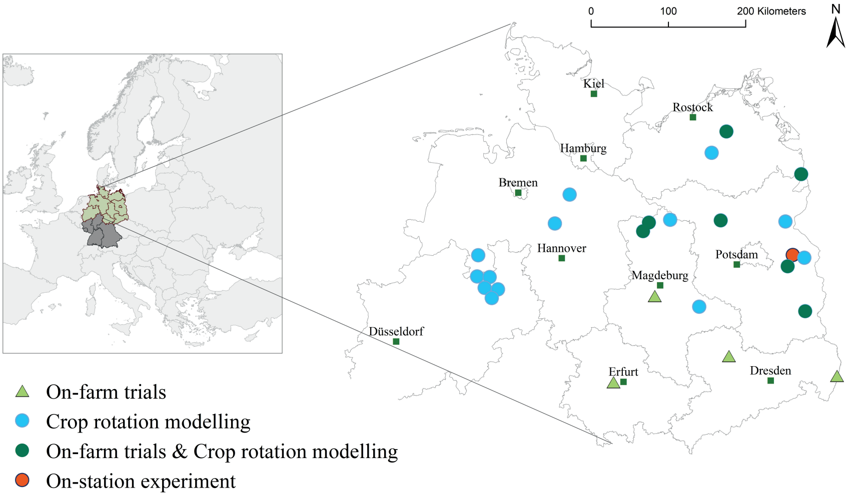

Figure 5. Map of the case study area in northern Germany (Paper V). Location of the farms participating in the on-farm research and crop rotation modelling, and the location of the on station experiment.

3.4.2 The DEED research cycle

We used the DEED research cycle (Giller *et al.*, 2015) for the co-design with farmers in this study following four steps:

1 Describe production constraints and opportunities with grain legumes: A semi-quantitative survey was conducted among 23 organic farmers covering current grain legume production constraints and opportunities and collecting data from 37 fields for the crop rotation modelling. In addition, farmers, advisors and scientists discussed strategies to overcome constraints with reference to the experiments and trials during four field days.

2 Explain the impacts of current grain legume rotations: Crop rotation modelling with ROTOR was used to assess the current weed infestation

risk, and the N and soil carbon balances. These indicators were selected because they were mentioned as important in the survey.

3 Explore alternative practices through testing technical options at the field scale: Because of the specific farmers constraints the tested options focused on weed management, tillage, fertilization, comparing cultivars, the effect of irrigation and pre-crop effects. A combination of on-farm trials and on-station experiments was implemented. The latter was analysed statistically and both were evaluated with farmers and advisors.

4 Re-design grain legume cropping systems: Workshops were held to identify strategies evaluated as 'successful' by single farmers and advisors, and to co-design new grain legume CS based on the results from the testing during step 3.

4 Results

4.1 Impacts of legumes at the cropping system scale

The developed framework allows a systematic evaluation of the impacts of CS taking rotational effects into account. In five case study regions, large numbers of CS were generated and compared to quantify differences in environmental and economic impacts between systems. On average, CS with legumes reduced N_2O emissions by 18 % and 33 % and N fertilizer use by 24 % and 38 % in arable and forage systems, respectively, compared to systems without legumes (Table 2). Nitrate leaching was similar with and without legumes in arable systems and reduced by 22 %, on average, in forage systems. On average gross margins were lower in arable systems compared to systems without legumes, while gross margins were larger in mixed systems (Table 2).

Table 2. Environmental and economic impacts (relative mean effects and range in %) of introducing legumes into arable and mixed cropping systems in five case studies across Europe

Indicators	Arable systems		Mixed systems	
	Mean effect (%)	Range (%) [a]	Mean effect (%)	Range (%) [a]
N_2O emissions	-18	-12 to -30	-33	-23 to -52
N-fertilizer use	-24	-17 to -40	-38	-27 to -58
Nitrate-N leaching	-7	-24 to +3	-22	-50 to +5
Gross margins	-14	-73 to +29	+21	0 to +62

[a] Range of differences across the case study regions.

The assessment identified significant trade-offs between environmental and economic impacts in arable systems e.g. nitrous-oxide emissions and gross margins (see Figure 6, as an example for arable systems in Brandenburg). In mixed systems with forage legumes, win-win situations were identified.

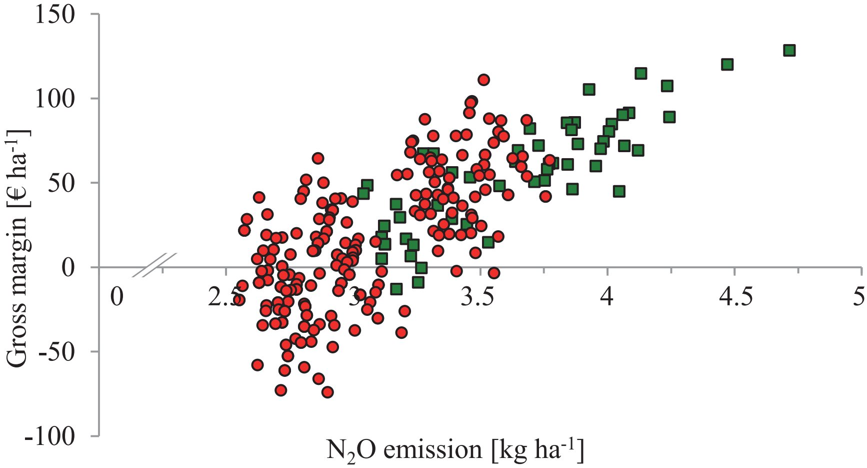

Figure 6. Gross margin plotted against N_2O emissions for arable cropping systems with (red circles) and without legumes (green squares) in Brandenburg (Paper I).

The impacts of legumes on economic and environmental services were different in each case study region and the framework allowed the comparison of specific CS. The 'current farming' system tended to have larger trade-offs than the majority of the generated systems with higher emissions and relatively high gross margins (see Figure 7 with an example for Brandenburg and the spider charts in Paper II for all other regions). The 'economic-environmental' optimum system with legumes (see section 3.2.4 for the selection criteria) significantly reduced trade-offs. On average across the five case study regions, the 'economic-environmental' optimum system with legumes had 32 % lower environmental impacts (combining all N-related impacts) and 21 % higher gross margins compared to 'current farming'. It was only in Brandenburg and Calabria that the gross margins were reduced with legume-supported arable systems, by 50 % and 25 % compared to 'current farming', respectively. The optimized CS without legumes also performed better than 'current farming' for many indicators. Compared to the optimized CS without legumes, the system with legumes reduced environmental impacts by 22 % and increased gross margins by 14 % on average across the five case study regions.

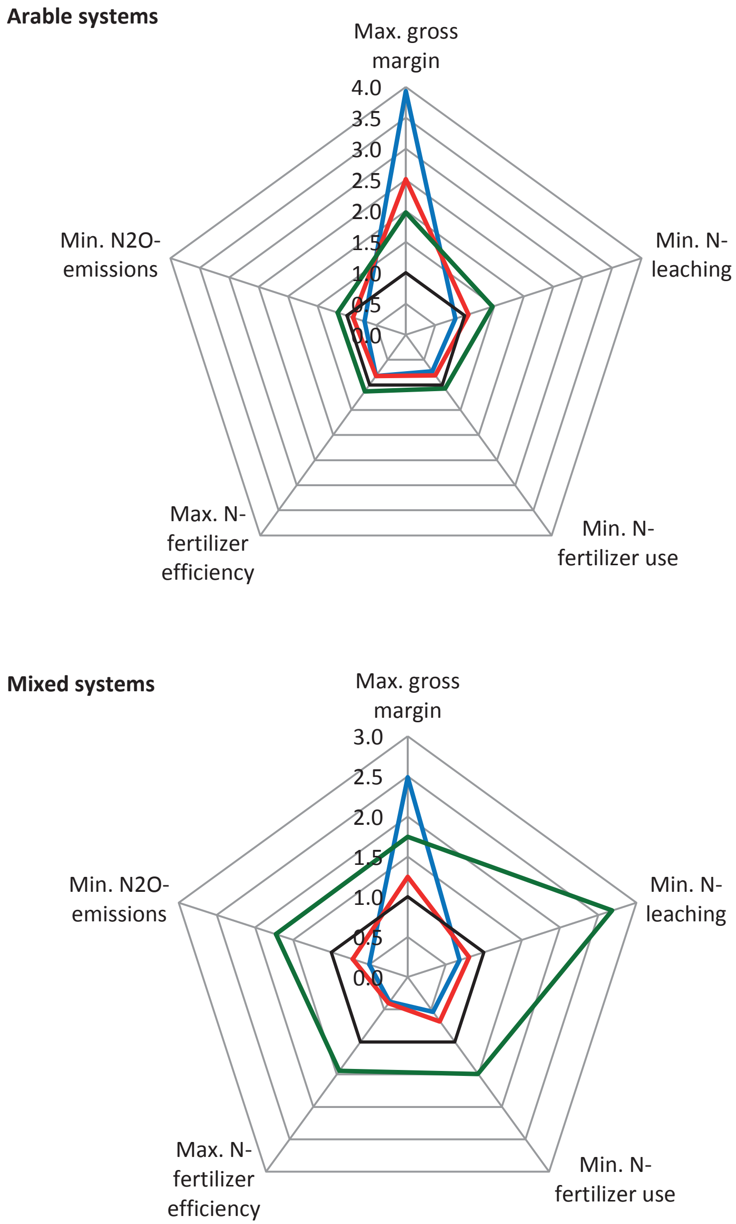

Figure 7. Multi-criteria assessment of arable and mixed cropping systems in Brandenburg (Paper II). Comparison of current farming (blue line), economic-environmental optimum system with legumes (green line) and without legumes (red line), and the mean impact (black line).Values are the ratio of the single impact relative to the average impact calculated for that indicator across all CS per region (outside values represent positive impacts).

4.2 Yield stability analyses

When the CV as the standard yield stability indicator was tested for the presence of a TPL-relationship is was found to be negatively and significantly correlated with yield in all three data sets used in this thesis: (i) the global yield data for wheat (adjusted $R^2 = 0.19$, $P < 0.001$), (ii) the global yield data for rye (adjusted $R^2 = 0.09$, $P < 0.001$), and (iii) the yield data from five LTEs (adjusted $R^2 = 0.21$, $P < 0.001$) (see Figure 8 A as an example). There was a highly significant linear relationship between log(mean) and log(variance) in all three data sets and in all cases TPL was valid, with regression slopes *b* being >1, but <2 (see Paper III and IV).

While the CV was dependent on mean yield, the aCV obtained by applying TPL was independent of yield in all tested data sets (see Figure 8 B as an example), showing that this new indicator can be used to estimate yield stability independent of the mean for different applications.

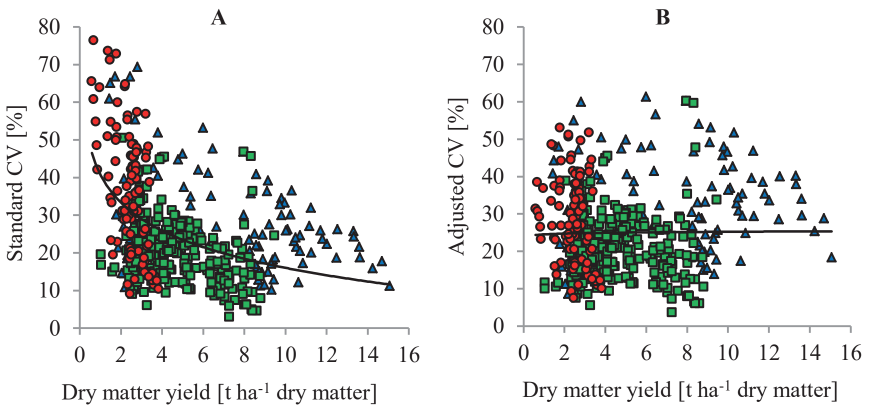

Figure 8. Relationship between yield and (A) the standard coefficient of variation (CV) and (B) the adjusted CV (Paper IV). Each data point represents the mean and variance of an 8-year period from long-term experiments for grain legumes (n = 100, red circles), broad-leaved crops (n = 96, blue triangles) and cereals (n = 275, green squares). The relationship is shown with a logarithmic regression line over all groups of crops (n=471).

A global trend of decreasing cereal yield stability over time was detected by applying the adjusted CV and this trend was captured less for rye and not at all for wheat with the standard CV (Table 3). The aCV was relevant because this dataset is characterised by large yield differences among crops and a significant increase in mean yield over the 50 years studied (Table 3).

Table 3. Relationship between time (decade), yield stability (CV and aCV) and mean yield of wheat and rye from a global data set over the period 1964-2013

	Wheat			Rye		
Relationhsip	Estimate	SE	P-value	Estimate	SE	P-value
Standard CV ~ decade	-0.009	0.023	0.694	0.078	0.027	<0.01
Adjusted CV ~ decade	0.066	0.018	<0.001	0.100	0.026	<0.001
mean ~ decade	0.041	0.002	<0.001	0.029	0.003	<0.001

Yield instability estimated with the aCV of grain legumes (30 %) was higher (P < 0.001) than that of autumn-sown cereals (19 %), but lower (P < 0.001) than that of other spring-sown broad-leaved crops (35 %), and only slightly greater (P = 0.042) than that of spring-sown cereals (27 %) (Figure 9). With the scale-adjusted CV, yield stability was estimated 9 % points higher for grain legume crops compared to an assessment with the standard CV.

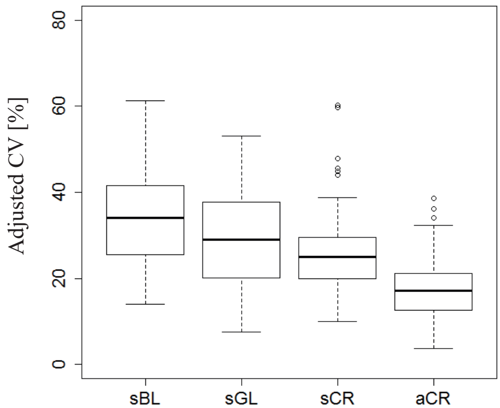

Figure 9. Yield stability of different crop groups estimated with the adjusted coefficient of variation (Paper IV). Comparison between spring-sown broad-leaved crops (sBL) (n = 75), spring-sown grain legumes (sGL) (n = 100), spring-sown cereals (sCR) (n = 117) and autumn-sown cereals (aCR) (n = 158). In each boxplot, the median is the black bar, the box covers the interquartile range, the whiskers cover the entire range of data, and circles indicate potential outliers.

Overall, using the aCV, yields of autumn-sown crops were 10 % points more stable than spring-sown crops (P < 0.001) and the grain legume yields were more stable than yields of other broad-leaved crops (P = 0.013). There was no difference in yield stability between grain legumes (all spring sown) and all non-legume spring-sown crops (spring-sown cereals and broad-leaved crops) (P = 0.845).

4.3 Re-design of legume-supported cropping systems

Cropping systems were re-designed in co-design with farmers, advisors and scientists by identifying specific practices that were tailored to specific farming contexts. Each practice that was a new option to manage weeds, increase grain legume yields or reduce soil tillage was the result of a detailed analysis of the constraints and opportunities (Describe phase), the identification of the impacts of current grain legume rotations (Explain phase), the tested alternative practices (Explore phase) and the re-design of CS (Design phase). An example of this co-design process is 'the development of soybean production' as a new crop (illustrated in Figure 10).

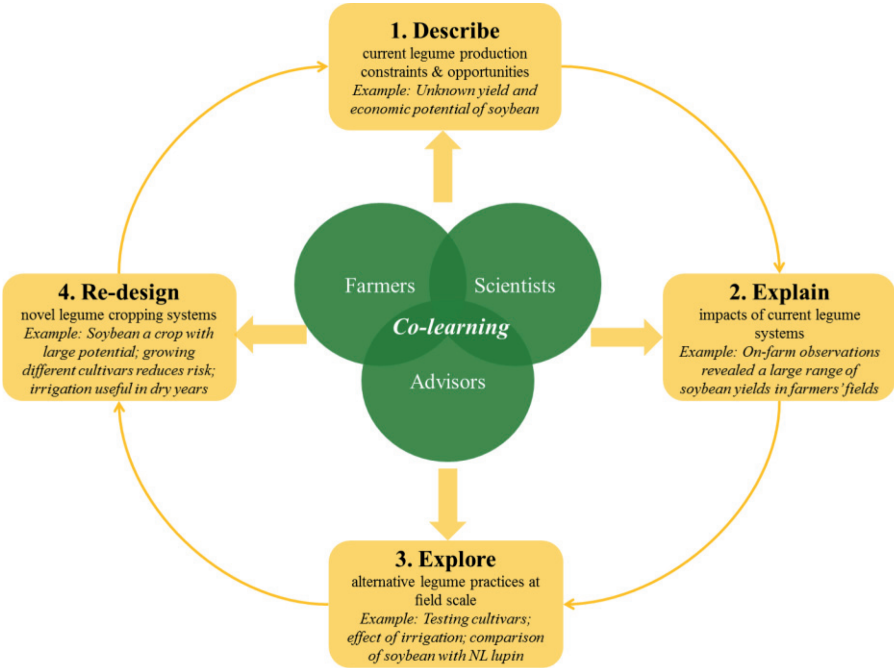

Figure 10. DEED framework for re-designing cropping systems (Paper V). The co-learning process with farmers, advisors and scientists, with an example for the development of soybean production.

The workshop at the start of the research cycle showed that farmers have little experience with soybean production, and little knowledge of what yields and gross margins could be expected compared to other crops. While high yields of soybean of up to 3.2 t ha^{-1} were possible on farmers' fields, they were not always achieved and did not always lead to sufficient revenues. Therefore, alternative cultivars for different markets and irrigation as means to stabilize yields were explored in on-farm trials and on-station field experiments as possible alternatives.

The trials and experiments revealed that there were advantages of growing soybean cultivars for both feed and food markets in the same year rather than for just one of these markets, because higher prices are paid for food-grade soybean but the yields for the food cultivars are often lower and do not always lead to higher gross margins. Irrigation increased yields significantly in dry years (2015-2016) but not in wet years (2014 and 2017). The comparison of soybean with NL lupin revealed larger average grain and protein yields for soybean due to pest and weed problems in NL lupin reducing the grain yield and a significantly higher protein content in soybean.

In the final step, another workshop and discussions on field days were used to re-design CS. Soybean was identified as a novel crop with large agronomic and economic potential (at least for organic farming). It was shown that growing soybean cultivars for food and feed would reduce risk, and that rainfed soybean cultivation is possible, but that flexible irrigation in dry years is a useful strategy to increase yields.

In total, nine practices were identified as successful alternatives by scientists in a mutual assessment with farmers and advisors, according to the results from the on-farm trials, on-station field experiments, and the recommendations and feedback from farmers and advisors during the DEED cycle.

The identified practices were:

1 Hoeing between rows in soybean with wider rows was a strategy evaluated to be successful by farmers and scientists to reduce weeds, in comparison to narrow rows without hoeing. This was successfully tested for NL lupin.
2 Direct seeding into crimped winter rye was a strategy for soybean that was possible only when water availability was sufficient.
3 Inoculation of soybean increased the amount of BNF and yield, but inoculation of NL lupin did not provide any advantages.
4 Cover crops after NL lupin reduced potential nitrate leaching effectively. In soybean, leaching risk was large after harvest but harvesting dates were too late to establish cover crops. Hence, strategies to reduce leaching following soybean e.g. direct sowing of a winter crop or undersowing of cover crops still need to be developed.
5 Reduced tillage in spring with the cultivator instead of the plough allowed more flexibility for field operations, timely sowing of soybean and potentially less energy use.
6 Soybean cultivation as a new crop achieved relatively high yields and large gross margins in organic systems.

7 Growing different soybean cultivars for food and feed markets was found to reduce risk. It requires good knowledge of the agronomic parameters of different cultivars and continued testing.
8 Irrigation increased soybean and NL lupin yields during dry years and should be used when possible during flowering and pod-filling in soybean. Rainfed cultivation of soybean provided sufficient yields especially in wet years.
9 NL lupin followed by a cover crop enhanced the growth of a subsequent crop due to the preserving and delayed release of N.

The practices identified are a starting point for further testing and adaptation by farmers.

5 Discussion and conclusion

5.1 Impacts of integrating legumes into cropping systems

5.1.1 The cropping system assessment framework

In contrast to other CS models that generate rotations such as ROTOR (Bachinger & Zander, 2007) and PRACT (Naudin *et al.*, 2015), the framework in this thesis generated rotations and evaluated crop production activities as two separate steps, before merging them to CS. The models mentioned above can produce only CS that are based on the fixed set of available crop production activities defined in the model database. The separation of the two steps in the framework allowed the generation of novel CS that were not previously considered. A particular strength is that it involves stakeholders from research and practical farming in the process of designing and assessing CS in order to fulfil both economic and environmental aims. This has the advantage that it can be done without the considerable expense of several years of testing many systems on farms or experiment stations across a wide range of environments.

For the generation of crop rotations, the philosophy was to follow the logic of ‘agronomist's thinking’ by making their knowledge on crop rotation design explicit. This could be utilized through formalizing crop rotation rules as a basis for the generation of ‘agronomically sound’ rotations. Agronomists tended to define very strict crop sequence and frequency restrictions, e.g., allowing only a small share of cereals that often did not allow the generation of current widely used rotations. Hence, several iterations were needed with agronomists.

For the evaluation of crop production, agronomists used information from regional statistics and estimated pre-crop effects to define the crop production activities. The strength of using expert estimations was that the pre-crop effects on yields consider factors such as root health, pests, weeds and disease that are often not considered in simulation models (Bergez *et al.*, 2010; Kollas *et al.*, 2015). While a number of insect, disease and weed dynamic models have been developed (Colbach *et al.*, 2017; Jones *et al.*, 2017), only some were coupled with crop models, and those were specifically for cotton and soybean (Jones *et al.*, 2017). The quality of the model outputs from the framework was dependent to a large extent on the quality of the expert knowledge, as is usually the case for rule-based models (Bachinger & Zander, 2007; Naudin *et al.*, 2015). Complementary to the evaluation by experts, the outputs of this static approach were compared with measured data for specific impacts, such as N leaching, and could be compared against the outputs of dynamic models in future.

Another approach to evaluate CS is the use of a life cycle assessment. While studies have been carried out to quantify the environmental impacts of introducing grain legumes into European crop rotations (Nemecek *et al.*, 2008; Köpke & Nemecek, 2010; Knudsen *et al.*, 2014) these work with a limited set of pre-defined rotations and do not include agronomic aspects such as the weed infestation.

A major limitation of the static framework is that soil-crop processes are considered only on an annual basis, which does not allow the identification of constraints during specific growth stages, such as growth limitations due to water deficit. Also the calculation of nitrate leaching and N_2O emission did not consider the variability of weather, yields and environmental impacts, and are used here to make only relative comparisons between CS (Paper II). While dynamic models could potentially simulate these processes in detail, these can so far model only a limited number of rotations and are currently unable to simulate all relevant rotational effects (Lorenz *et al.*, 2013; Kollas *et al.*, 2015; Jones *et al.*, 2017). However, the static framework is also very sensitive to the estimation of pre-crop effects and requires a good combination of results from experiments, especially those reported in meta-analyses (Angus *et al.*, 2015; Preissel *et al.*, 2015; Cernay *et al.*, 2017; Franke *et al.*, 2018) and expert knowledge for the specific situations where the framework is applied.

The framework presented here could be used to compare other CS and address new research objectives, such as to reduce yield stability. An indicator for yield stability could be developed to assess CS and integrated into the framework. This ‘cropping system yield stability indicator’ could be utilising the methods developed in paper III. The indicator could be calculated as the mean aCV over a CS in order to allow comparisons between CS with different lengths. Yield stability should also be considered in the economic calculations, i.e. the variation in yields and prices and be reflected in the gross margin (LMC International, 2009). A challenge would be the availability of regional specific yield data and prices that is also affected by scaling and aggregation. Other important possible applications are to explore further sustainability indicators such as soil carbon sequestration and biodiversity.

5.1.2 Environmental and economic impacts

The major environmental effect when grain and forage legumes were integrated into CS, was related to N input and N losses. The reduction in the use of N fertilizer was 17-40 % in arable and 27-58 % in forage systems and was mainly attributable to the N added through the BNF of legume crops (Iannetta *et al.*, 2016). The greatest savings were made in forage systems, because the perennial legumes included can fix up to 350 kg ha^{-1} (Carlsson & Huss-Danell, 2003), which is much more than the corresponding figures for the annual legumes, pea and faba bean, that fix about 130 kg ha^{-1} (Peoples *et al.*, 2009a).

Lower rates of N input also resulted in lower N_2O emissions in CS with legumes because of the direct relationship between N fertilizer input and N_2O emissions (Rees *et al.*, 2012; Buckingham *et al.*, 2014). N_2O emissions were on average 18 % and 33 % lower in arable and forage systems, respectively. Again, the differences were greater for forage systems than for arable systems due to the differences in N fertilizer use and larger amounts of BNF in forage systems. The assumption that 1 % of each kg of N fertilizer is released as field-based N_2O (IPCC, 2006) was used, and this is currently under consideration. Philibert *et al.* (2012) calculated lower emission factors (EF) when the amount of N applied was below 160 kg ha^{-1} and Hinton *et al.* (2015) estimated EF to be between 0.28 % and 1.35 % of applied N depending on the N input. Rees *et al.* (2012) concluded from a meta-analysis using measured data from Europe that annual emissions from arable sites were significantly greater than predicted by IPCC and mainly influenced by the N input. The EF used in the framework is clearly approximate, but it can only be improved in a dynamic approach, which is not compatible with the static framework. Goglio *et al.* (2018) found that the combination of the simple carbon model by Andrén and Kätterer (1997) with the IPPC method provides better results than the IPPC method alone.

Nitrate leaching was on average similar with and without legumes in arable CS because optimal management of N-rich residues of grain legumes and grass-clover mixtures was assumed to avoid large losses after harvest and termination. The risk of nitrate leaching is high on sandy soils (Askegaard *et* *al.*, 2005; Benoit *et al.*, 2014), and especially in systems that receive high amounts of fertilizers (Eriksen *et al.*, 2015). Since less N fertilizer was used in both arable and mixed systems with legumes (Table 2), leaching tended to be similar in arable and lower in mixed systems. Grain legumes were followed by winter cereals taking up some of the N left by the legumes. Nitrate leaching could be reduced in legume-supported systems by using cover crops in both arable and mixed systems (Askegaard *et al.*, 2005; Beaudoin *et al.*, 2005; Benoit *et al.*, 2014; Plaza-Bonilla *et al.*, 2015). Cover crops were not tested in the study presented in Papers I and II. In the framework, nitrate leaching was calculated in a relatively simple way from the N surplus and De Notaris *et al.* (2018) showed that, at the rotation level, N leaching was positively related to N surplus in a long-term crop rotation experiment in Denmark. However, they found that the overall effect of N input and surplus on N leaching was lower than the effect of use of cover crops. Therefore, the effect of CS with and without cover crops should be included in the framework if the aim is to reduce nitrate leaching.

The break-crop effect is another aspect that increases the efficiency of utilizing N after a legume. Roots of a given crop are healthier after an unrelated crop has been grown, because pathogen populations are reduced (Angus *et al.*, 2015), allowing more N to be taken up by the crop, and reducing the availability of N for leaching. Soil tillage could in some circumstances be reduced before and after grain legumes (Luetke-Entrup *et al.*, 2003; López-Bellido *et al.*, 2004), which could reduce losses of N and increase gross margins of legume-supported systems (Preissel *et al.*, 2015).

Economic performance is regarded as a key driver responsible for low adoption of grain and forage legumes in CS by farmers (Von Richthofen *et al.*, 2006a). Conversely, Luxemburgish farmers mentioned a lack of knowledge and extension services for these crops as the main barrier for not cultivating grain legumes, rather than economic issues (Zimmer *et al.*, 2016). This is not a contradiction, since good knowledge on agronomy and markets are needed before growing the crop. At the crop level, grain legumes often have lower gross margins than cereals and oilseed crops (Preissel *et al.*, 2017). Although substantial yield benefits from legumes to the following crops have been widely observed across the world (Angus *et al.*, 2015; Preissel *et al.*, 2015; Cernay *et al.*, 2017; Franke *et al.*, 2018) they are rarely considered in the economic evaluations of CS (Zander *et al.*, 2016; Meynard *et al.*, 2018). Paper II showed for the first time that the economic performance of grain and forage legumes was improved when assessed at the CS scale across a large range of case studies and CS. The difference in gross margin between systems with and without legumes were relatively small when all rotational effects were included in Paper II, ranging between -67 and +106 € ha^{-1} in arable systems with grain legumes and from 0 to +50 € ha^{-1} in mixed systems with forage legumes.

Thus, this thesis demonstrates the need to assess legumes at the CS scale and take the environmental and economic impacts in a combined way into account. It describes a framework that allows such an integrated assessment and comparing a large range of CS.

The framework with the integrated indicators advances the state of the art in assessing CS systematically, scientifically and in a standardized way over a range of biophysical and socio-economic conditions. The application has shown substantial positive environmental impacts of integrating legumes into CS. There is potential for further refining the used indicators and broadening the scope of impacts assessed by integrating additional indictors. The application also showed that in many systems, but not all, modest economic trade-offs occur. A yield stability indicator for the CS level could contribute another significant economic criterion.

5.2 Stability of grain legume yields

5.2.1 Grain legumes are as stable as other spring crops

Yields of grain legumes were shown to be less stable than those of winter-sown cereals but not significantly different from other spring-sown crops in LTEs (Paper IV). These novel findings contrast with previous research (Hawtin & Hebblethwaite, 1983; Peltonen-Sainio & Niemi, 2012; Cernay *et al.*, 2015) and with farmers' and experts' perceptions (Von Richthofen *et al.*, 2006a; EIP-AGRI, 2014; Zimmer *et al.*, 2016; PGRO, 2018). First, farmers may perceive grain legumes to be less stable because of relatively low market prices and the poor development of value chains (Preissel *et al.*, 2017; Meynard *et al.*, 2018). Although changes in grain legume producer prices followed the changes for soybean and wheat (LMC International, 2009), European pea and faba bean producer prices did not fully follow the price increases of soy-based feed ingredients. Agronomic constraints with pests, diseases and weeds in grain legumes (Watson *et al.*, 2017) might be higher at the farm level and are less visible in LTEs. Second, all previous studies have used indicators that do not correct for the association of the variance with the mean yield, which over-emphasizes variation in low-yielding crops as shown in the results of Paper IV. Third, when yield data are aggregated at the national level in official statistics, there is a tendency that the apparent yield stability increases with the size of the harvested area (Cernay *et al.*, 2015) (see section 5.2.2 on the effects of scaling). The latter two aspects could have resulted in lower apparent stability for grain legumes with generally lower yields and that are grown on smaller areas than many other crops. This thesis described and applied the aCV with field-level data from LTEs, where all crops are grown in the same proportion and on the same plot size. Thereby, the study described in this thesis dealt effectively with the inappropriate effects of scale on measures of stability that have been encountered in earlier studies.

Still, yield stability needs to be increased to make grain legumes more compatible. Besides long-term breeding efforts (Rubiales *et al.*, 2015), farmers need direct (short-term) strategies to deal with yield instability. In this thesis, management strategies were identified to reduce yield instability, to increase yield and to provide other benefits that make grain legume production more attractive to farmers (Paper V). Growing soybean and NL lupin in a wider row spacing and hoeing between rows were effective ways to reduce weed infestation and similar effects have been shown in spring cereals (Melander *et al.*, 2018). Inoculation of soybean seed increased BNF and yield and irrigation increased yield in dry years. If implemented over longer time periods these measures will potentially lead to higher yield stability.

While the findings from the Papers III and IV advance our understanding of yield stability assessments and the effects of scale, they also provide some practical messages.

- Spring crops are generally less stable than winter crops. Comparisons of grain legumes with winter crops are therefore less meaningful.
- Currently grain legumes are often grown on less fertile soils than other crops, therefore comparisons with crops grown on better soils are less meaningful and similar conditions have to be ensured.
- The CV should be used as a relative yield stability indicator only when crop yields are relatively similar.
- If differences in yield are large, the CV needs to be adjusted. A simple guide and example calculation for the aCV is provided by Döring and Reckling (2018) in addition to Paper III.

5.2.2 Scaling effects in the stability indicator and yield data

The scale (or aggregation) of the yield data as well as how the yield stability indicator considers scaling issues in crop yields (large differences in mean yield) plays a key role for yield stability assessments.

The larger the scale of yield aggregation, the more likely the effects of factors affecting crop yield may be averaged out, resulting in lower variability (Rudstrom *et al.*, 2002; Popp *et al.*, 2005). The results from this thesis demonstrated that there is a difference between the CV obtained from field-level data and that obtained from aggregated data where variability has been drastically reduced. The average CV value of 26 % from the LTEs (Paper IV) was twice as large as the average CV of 12 % (Peltonen-Sainio & Niemi, 2012) and 13 % (Cernay *et al.*, 2015) in the studies with national aggregated yield data. The values are much greater than global patterns of modelled crop yield stability with CV values of the ensemble median ranging between 2.3 % and 8.3 % for rice, maize, soybean and wheat (Müller *et al.*, 2018). Scaling methods should therefore be used carefully to avoid complex multi-scale problems (Ewert *et al.*, 2011).

Long-term experiments have not previously been used sufficiently for assessment of yield stability and this resource could be exploited more effectively (Johnston & Poulton, 2018). There are several hundred LTEs available worldwide (620 are listed in a global assessment by Debreczeni and Körschens (2003) alone) and there are several national platforms that provide access to the meta-data. A network of LTEs, as proposed by Stützel *et al.* (2016), could help to better utilize the existing resources. Nevertheless, there are at least three major problems related to the use of data from LTEs. First, not all crops of current economic importance are grown; second, many experiments have unbalanced designs so that not all crops are grown in the same year; and third, the small plot size can result in lower stability of crops that are disadvantaged in small plots (Rebetzke *et al.*, 2014). The latter is due to yield losses because of biotic stresses, as is the case for grain legumes. Kravchenko *et al.* (2017) found that the yield gap between experiments with small plots and field-scale experiments with large fields was more pronounced for soybean and maize than for wheat. This suggests that use of large plots would remove a structural bias against grain legumes and other crops that are sensitive to competition and other biotic stresses.

While this scaling refers to the aggregation of data over an increasing area, the yield stability indicator also needs to account for scaling that is related to the size of the mean. With increasing mean there are effects on other statistical parameters, i.e. they vary depending on how large the mean is (Döring *et al.*, 2015). The coefficient of variation (CV) is often calculated in ecological research, which divides the variability (standard deviation) across years by the mean yield over the same period to obtain a 'relative yield stability' indicator (Knapp & van der Heijden, 2018). However, the results presented in Papers III and IV show significant relationships between yield and the standard CV in three independent data sets. This relationship explains why stability by default increases with yield but this is rarely considered even in very recent studies (Raseduzzaman & Jensen, 2017; Hoffmann *et al.*, 2018; Knapp & van der

Heijden, 2018; Müller *et al.*, 2018). Other existing variance- and regression-based indicators used to analyse yield stability (Piepho, 1998) do not account for this dependence on mean yield. Thus, the aCV is an important addition to the set of stability indicators available in plant sciences, such as described by Piepho (1998), Ferreira *et al.* (2006) and Dehghani *et al.* (2008).

As shown in Paper III, scaling is also relevant when estimating changes in yield stability over time. Applying the aCV showed that yield stability of wheat and rye significantly decreased globally over the last 50 years, which was not detected with the standard CV. In accordance with this, Reckling *et al.* (2018) found indications that yield stability of grain legumes and other crops decreased between 1953 and 2015 in two LTEs in Germany and Sweden. This decrease in stability may be associated with an increase in climate variability and a lack of adaption to these changes (Ray *et al.*, 2015; Tigchelaar *et al.*, 2018). Up to now, there is a lack of research on the impacts of climate variability on grain legume yield stability. It is necessary to quantify how temporal yield stability changes over time for different grain legume species and across environments. This could benefit from a combination of statistical analyses and simulation modelling, such as used by Lobell *et al.* (2011a) and Hochman *et al.* (2017), based on data from LTEs across different environmental conditions taking scaling effects into account.

This thesis advances the ability to study yield stability while avoiding scale-dependent biases through a new indicator and through highlighting the impact of the data used. Through the improved method, differences between CS can be assessed and long-term trends, including impacts of climatic change and agricultural developments, can be investigated.

5.3 Re-designing cropping systems

5.3.1 Involving farmers in the re-design of cropping systems

The results from Paper V provide new insights in how to re-design CS with grain legumes using a participatory approach. It applied the DEED research cycle with large-scale farmers in Europe, that has so far mainly been used with smallholder farmers and stakeholders in Sub-Saharan Africa (Giller *et al.*, 2011; Descheemaeker *et al.*, 2016; Falconnier *et al.*, 2017). Following the cycle helped to focus on farmers' perceived production constraints and opportunities, i.e., low yield stability and the risk of high weed infestation. The perception of low yield stability was also found to be very important in earlier surveys among farmers in Europe (Von Richthofen *et al.*, 2006a; Zimmer *et* *al.*, 2016). The risk of high weed infestation is also recognized in grain legumes (Avola *et al.*, 2008; Rubiales & Fernández-Aparicio, 2012; Döring, 2015), but farmers still lack appropriate strategies to deal with this risk.

This thesis demonstrates the need for complementing formal knowledge from on-station research and modelling with on-farm trials and local knowledge, to effectively re-design farmers' CS. Farmers' knowledge on production constraints and opportunities to improve the cultivation of NL lupin (the existing crop) and explore options on how to cultivate soybean as a novel crop was central in the study described in Paper V. The co-design process provided new insights into the use of different methods for the re-design of CS with legumes and identified tailored options for specific farm contexts (see the nine practices in section 4.3). The on-farm trials were managed differently by each farmer with the aim of designing innovative farming practices (Falconnier *et al.*, 2017; Catalogna *et al.*, 2018) rather than comparing treatments statistically as would be done for variety testing (Schmidt *et al.*, 2018). The practices tested were partly novel such as the direct sowing of soybean into a crimped winter rye cover crop but also classical such as inoculation, weed control and irrigation. A similar co-design process to generate a relevant basket of options for climbing bean cultivation was taken with a diversity of smallholder farmers in Uganda (Ronner, 2018). That study aimed to distinguish preferences of farmers of different gender and socio-economic backgrounds for different technologies, but could not identify consistent recommendations about the suitability of technologies for different types of farmers. Similar to the results in Paper V, it highlights the importance of a basket of options with flexible combinations of practices for farmers' individual selection. The practices identified in Paper V should therefore be seen as flexible alternatives to current farming that could be adapted further to suit the specific needs of farmers, rather than as fixed technology packages taken from formal research of options (Sumberg *et al.*, 2003). While most studies on design in agronomy have focused on the invention part of the design process, Prost *et al.* (2018) emphasize the importance of examining the way in which the designed options are implemented by farmers and of continuing the design process over time. This is also a shortcoming of the study presented in this thesis.

Another option for the design of CS with stakeholders is the use of multi-criteria analysis such as done by Pelzer *et al.* (2017). They identified innovative legume-supported CS with higher sustainability impacts than current CS. Nevertheless, in their study, CS were designed without an active involvement of farmers. In the study presented in Paper V, the focus was on the re-design with an active contribution by farmers with less emphasis on the assessment. Using MASC, the multi-criteria analysis tool used by Pelzer *et al.*

(2017) could be a useful addition to compare the designed CS in this thesis in terms of their sustainability impacts.

5.3.2 The role of different methods for cropping systems design

In Paper V a combination of methods was used in the interaction with farmers, which differ in their suitability for working with farmers and provide different contributions to the assessment and design process of CS.

Farmers valued the crop rotation modelling with ROTOR as a relevant and useful instrument to highlight problems of the production system such as weed infestation, nutrient deficiency and soil organic carbon losses. Other tools are available for other contexts (Naudin *et al.*, 2015; Colbach *et al.*, 2017; Topp *et al.*, 2017). Models simultaneously consider different indicators (N, soil organic carbon, weeds etc.) that are seldom all measured in the field and allow multi-criteria analysis to assess the feasibility of new systems (Pelzer *et al.*, 2017). In Paper V, models were seen as a tool for the diagnosis of the system and its components, to provide useful input for discussions about the farm management and allow an *ex-ante* assessments of adaptation options as described by Topp *et al.* (2017). Most models are still relatively 'user unfriendly' and the results are difficult to interpret for farmers (Jones *et al.*, 2017). Thus, simplified tools are needed that are freely available, easy to understand and ready to use for different farm and production systems. In ideal cases, default values are included along with options to modify these for advanced users. Examples of such tools are: NDICEA (van der Burgt *et al.*, 2006), FarmDESIGN (Groot *et al.*, 2012) and ROTOR (Bachinger & Zander, 2007).

On-farm trials are more meaningful for practical farming than on-station experiments, they are more flexible and involve farmers from the design phase until the interpretation of the findings, so they can result in practical conclusions (Falconnier *et al.*, 2017). They are especially important for assessing how new technologies will function under practical farm conditions (Kravchenko *et al.*, 2017). However, on-farm trials are less precise than on-station ones (Schmidt *et al.*, 2018) and the results might be very meaningful for one farm but not for others. This reduces the possibility for more farmers adopt an innovation within a particular region (scaling up) and to transfer an innovation generated in one region to another region (scaling out) (Sinclair, 2017). Therefore, the number of on-farm trials should be sufficiently large depending on the research questions. Although on-station experiments are valuable because of the accuracy of the management and reproducibility, obtained yields are generally higher than in practice due to the controlled conditions. The differences between on-station results and on-farm practice are not systematic and affect different crops to different degrees. For example, Kravchenko *et al.* (2017) found a larger yield gap between field experiments and on-farm trials for soybean and maize than for wheat because of problems with weed management in on-farm trials. Innovation is locally driven in on-farm trials in a bottom-up process (Sinclair, 2017), and on-station experiments and modelling compare innovations across locations and systems. If used complementarily, they offer scope for scaling-up and -out of innovations (Sinclair, 2017).

While Doré *et al.* (2011) highlighted on-farm research as a new avenue in agronomy, the results from this thesis suggest that the combination of methods is more valuable than using one method alone to effectively re-design CS. Also, by combining the methods, they support and re-inforce each other. On-farm trials generate new research questions to be tested with models and on-station, and models can test existing knowledge about CS. However, since experiments require more resources than modelling studies, there is a general tendency towards a decreasing number of publications reporting experimental results during recent decades as has been shown for Central Asia (Hamidov *et al.*, 2016). At the same time, modelling studies have increased in number (Hamidov *et al.*, 2016) although they depend on the data generated in field experiments for model calibration and validation (Rötter *et al.*, 2018).

A major limitation of the research process described above is that it is very time- and resource-intensive so it has limited applicability for a regular 3-year research project. Following the DEED research cycle requires time for the assessment, analysis and testing of new systems and possibly a second or third round of the cycle. The approach might be more suitable for large research projects (Sinclair, 2017) or demonstration networks involving many farmers and other actors. In such projects, the described approach can provide guidance for identifying innovative systems and for the scaling-up and -out of the results (Sinclair, 2017).

5.4 Challenges and opportunities for legumes in Europe

It is easier for farmers to capitalize on the opportunities for forage legumes than on those for grain legumes because the forages (i) offer better economic and environmental performance than the grain legumes, (ii) have a high feed value for livestock and (iii) are relatively simply integrated into existing temporary grassland. The results from Paper II showed win-win situations where the cultivation of forage legumes is economically attractive and increases environmental benefits. Besides the impacts assessed in this thesis, the perennial nature of forage legumes offers opportunities for improved biodiversity (Stein-Bachinger & Fuchs, 2012) and soil organic carbon content (Jensen *et al.*, 2012), lowered risk of both soil erosion (Jensen *et al.*, 2012) and weed infestation (Håkansson, 2003). Important constraints with forage legumes include crop establishment, the maintenance of an appropriate balance (about one third) of legumes in mixtures with grasses to achieve the maximum benefits (Suter *et al.*, 2015), lower persistence than grass under grazing (Phelan *et al.*, 2015), high risk of livestock bloat from some species (Dewhurst *et al.*, 2009) and difficulty of conservation as silage or hay (Phelan *et al.*, 2015). The adoption of forage legumes is mainly restricted to mixed farms with crop and livestock production. The specialization of farming and associated spatial decoupling of livestock and crop production is probably a major reason for the low proportion of forage legumes in Europe (Lemaire *et al.*, 2015). Alternative avenues for forage legume production by arable farms are the collaboration with livestock farmers and the exchange of materials such as forage, grain, straw and manure (Lemaire *et al.*, 2015; Martin *et al.*, 2016; Asai *et al.*, 2018). Other options are to deliver forage legumes as green biomass to biogas plants (Tidåker *et al.*, 2014) and to biorefineries (Papendiek *et al.*, 2016; Parajuli *et al.*, 2017), and to produce dried fodder especially from alfalfa (*Medicago sativa* L.).

There are situations where introducing grain legumes into arable systems will give both economic and environmental advantages (Paper II). This is where grain legumes achieve high prices for human food use, e.g. in Romania for common bean, and where grain yields are relatively high as in the UK for faba bean. These benefits are currently not reflected in farmers' rotations because value chains are not sufficiently established (Meynard *et al.*, 2018) and other forces favour specialization on cereal crops (Magrini *et al.*, 2016). Therefore, grain legumes are currently not economically competitive with the small-grain cereals in most cases (Paper II). Farmers are not rewarded for the positive environmental services which legumes deliver (Zander *et al.*, 2016). In Europe, the number of food products containing grain legumes such as beans, lentils or soybeans have increased by 39 % between 2013 and 2017 (TRUE, 2018). While this trend will increase the demand for European-grown grain legumes to some extent along with awareness among consumers (Profeta & Hamm, 2019), it is unlikely to lead to large increases in production areas since the market is still relatively small. A bigger increase in demand and effect on legume production can be expected when European grain legumes are used in feed value chains or in the non-food sector.

In order to utilize the opportunities associated with grain and forage legumes, agronomic practices need to be further develop in both grain and forage legumes, benefitting from the results in Paper V. Besides the agronomy, regional supply chains need to be developed for feed, forage, food and non-food that could increase the competitiveness of legumes in Europe (Voisin *et al.*, 2014; Meynard *et al.*, 2018). Examples are GMO-free products, novel foods based on plant proteins and other high added-value niche markets (Voisin *et al.*, 2014; TRUE, 2018). The results from Paper II and other studies such as by Pelzer *et al.* (2017) and Lötjönen and Ollikainen (2017) contribute to making the services of grain and forage legumes explicit. This is the basis so that they could be considered in farmers' economic calculations (Zander *et al.*, 2016). Policy can support grain legumes in cases where markets are not yet established or where the services that are relevant for society are not valued. Policy can provide direct subsidies or indirect support such as the use of legumes in ecological focus areas in the common agricultural policy. Pros and cons of different policy options were compared by Bues *et al.* (2013). According to Magrini *et al.* (2018) strong support is required from public institutions to coordinate the transition of the agrifood sector to include legumes for reasons of human health and environmental sustainability.

5.1 Outlook

This thesis contributes to addressing the research need for an integrated assessment of legume-supported CS, evaluating the opportunities and challenges of legumes and identifying avenues for intensification of legume production in European agriculture. By developing the research approach of this thesis further, the assessment could be broadened, refined and integrated with a continued effort for designing CS that make better use of the ecosystem services of legumes by enhancing their production.

Future research should broaden the geographic coverage of the assessment with the CS framework to other regions across Europe. In areas where the CS framework has not yet been applied, such as in south-western Europe, the impacts of integrating legumes into CS could be different, as indicated for example by contrasting findings about higher carbon footprints and N leaching in south-western France (Plaza-Bonilla *et al.*, 2018). The framework could be further modified to integrate cover crops along with additional environmental indicators such as carbon sequestration and yield stability.

The yield stability indicator developed and tested in this thesis provides insights into interspecific differences in yield stability and could contribute to the design of more stable systems. To this end, crop yield stability should be assessed in other systems, e.g. production of winter-season legumes in Mediterranean and Atlantic regions, because it was found that winter crops were significantly more stable than spring crops in northern Europe. The effect of the lower yield stability of grain legumes and other spring crops on the average yield stability of CS is another research question to investigate. Beyond the research questions addressed in this thesis, there is a need to study the causes and main drivers for yield stability in grain legumes. Since results on cereals (Paper III) and grain legumes from two LTEs in Sweden and Germany (Reckling *et al.*, 2018) showed that temporal yield stability decreased over time, future research should investigate the effect of climate variability on yield stability and to identify practical adaptation strategies to design more resilient CS.

The thesis identified avenues for enhancing legume production for a specific study case, where soybean has shown to be a useful crop in a cool temperate climate. For soybean to become a major crop in northern Europe, there is a need to identify and potentially breed adapted both feed-grade and food-grade cultivars. Since soybean leaves much residual nitrogen in the soil, agronomic practices ought to be developed that reduce the risk of post-harvest nitrate leaching. Cover crop solutions to avoid leaching need to be developed that would suit the late harvesting time of this crop. Future research should also identify optimal subsequent crops that can take up much of the residual N and utilize the positive pre-crop effect. Yield stability of soybean should be assessed because this crop might be more stable than other grain legumes due to better tolerance to drought and fewer problems with pests and diseases. On-farm trials across a gradient of different soil and climatic conditions would be useful to identify the conditions under which soybean can be grown and how it can be integrated into CS.

Since research projects are often too short to involve farmers actively and implement on-farm trials, farmer-scientist networks need to be established that have a longer time frame than projects. In these networks, different objectives can be addressed, such as identifying CS for increasing resource-use efficiency. Since avenues for enhancing legume production are specific to local and individual conditions, a substantial improvement of legume production across regions or even Europe-wide requires many local-scale efforts that include not only further scientific research but also the involvement of extension services, practice networks and other stakeholders.

5.2 Conclusion

The aim of this thesis was to design and assess legume-supported CS in European agriculture. With newly developed and refined assessment methods, it was shown that integrating legumes into CS can provide substantial environmental benefits under different bio-physical and socio-economic conditions, while reducing income and yield stability modestly in some cases. The assessment methods and a case study with a co-learning process with farmers identified potentials for re-designing CS effectively.

First, a novel CS assessment framework was developed using a static and rule-based approach with the objective of considering crop rotations and rotational effects on environmental and economic indicators (Paper I). The framework enabled for the first time a systematic assessment of diverse CS and the effect of grain and forage legumes across a range of European regions, while taking the CS perspective. This perspective and the large range of CS assessed made the ecosystem services of legumes visible and brought out options for designing new systems.

Second, the economic and environmental effects of integrating legumes into CS were investigated with the objective of identifying the potentials and limitations of increasing legume cultivation in Europe (Paper II). Legumes reduced negative environmental impacts significantly in almost all test regions in arable and mixed systems. Trade-offs between economic and environmental impacts occurred when grain legumes were integrated into CS. Systems with forage legumes resulted in synergies by increasing economic benefits and reducing environmental impacts compared to systems without legumes.

Third, a novel method to quantify yield stability by adjusting the standard CV was developed with the objective of removing the dependence on the mean yield (Paper III). Accounting for Taylor's Power Law in a scale-adjusted CV removed the dependence on the mean yield in three data sets and enabled the scale-adjusted CV to detect a long-term yield stability decrease in cereals. The suggested method is an alternative to estimate yield stability more conclusively than the frequently applied indicators.

Fourth, yield stability of grain legumes was assessed with the objective of comparing it with other crop species using field-level data from LTEs from northern Europe (Paper IV). Applying the scale-adjusted CV showed that the yield stability of grain legumes and other spring crops is in a similar range, while winter cereals were significantly more stable.

The fifth objective was to re-design CS using a participatory approach based on farmers' perceived constraints and opportunities in growing legumes (Paper V). Working with farmers in a co-learning process highlighted the need to integrate formal knowledge from on-station experiments and modelling with on-farm research including the views of local farmers. The practices identified through the combination of methods increased the benefits of the most important services of grain legumes, namely provisioning of protein, N fixation and rotational effects, and reduced potential constraints of weeds and nitrate leaching. Implementing these practices will contribute to making the growing of grain legumes economically and environmentally more sustainable.

References

Andrén, O. & Kätterer, T. (1997). ICBM: The introductory carbon balance model for exploration of soil carbon balances *Ecological Applications,* 7(4), pp. 1226-1236.

Angus, J.F., Kirkegaard, J.A., Hunt, J.R., Ryan, M.H., Ohlander, L. & Peoples, M.B. (2015). Break crops and rotations for wheat. *Crop and Pasture Science,* 66(6), pp. 523-552.

Annicchiarico, P. (2002). *Genotype x Environment Interactions - Challenges and Opportunities for Plant Breeding and Cultivar Recommendations*. Rome: Food and agriculture organization of the united nations (FAO).

Asai, M., Moraine, M., Ryschawy, J., de Wit, J., Hoshide, A.K. & Martin, G. (2018). Critical factors for crop-livestock integration beyond the farm level: A cross-analysis of worldwide case studies. *Land Use Policy,* 73, pp. 184-194.

Askegaard, M., Olesen, J.E. & Kristensen, K. (2005). Nitrate leaching from organic arable crop rotations: effects of location, manure and catch crop. *Soil Use and Management,* 21(2), pp. 181-188.

Avola, G., Tuttobene, R., Gresta, F. & Abbate, V. (2008). Weed control strategies for grain legumes. *Agronomy for Sustainable Development,* 28(3), pp. 389-395.

Bachinger, J. & Zander, P. (2007). ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. *European Journal of Agronomy,* 26(2), pp. 130-143.

Beaudoin, N., Saad, J.K., Van Laethem, C., Machet, J.M., Maucorps, J. & Mary, B. (2005). Nitrate leaching in intensive agriculture in Northern France: Effect of farming practices, soils and crop rotations. *Agriculture, Ecosystems & Environment,* 111(1), pp. 292-310.

Becker, H.C. (1981). Correlations among some statistical measures of phenotypic stability. *Euphytica,* 30(3), pp. 835-840.

Becker, H.C. & Léon, J. (1988). Stability analysis in plant breeding. *Plant Breeding,* 101, pp. 1-23.

Benoit, M., Garnier, J., Anglade, J. & Billen, G. (2014). Nitrate leaching from organic and conventional arable crop farms in the Seine Basin (France). *Nutrient Cycling in Agroecosystems,* 100(3), pp. 285-299.

Bergez, J.E., Colbach, N., Crespo, O., Garcia, F., Jeuffroy, M.H., Justes, E., Loyce, C., Munier-Jolain, N. & Sadok, W. (2010). Designing crop management systems by simulation. *European Journal of Agronomy,* 32(1), pp. 3-9.

Bergkvist, G. & Båth, B. (2015). Nitrogen fertiliser dose influence the effect of two year rotational leys with grass or clover/grass on other crops in the rotation. *Aspects of Applied Biology,* 128, pp. 133-139.

Berthet, E.T., Hickey, G.M. & Klerkx, L. (2018). Opening design and innovation processes in agriculture: Insights from design and management sciences and future directions. *Agricultural Systems,* 165, pp. 111-115.

Bloch, R., Knierim, A., Häring, A.-M. & Bachinger, J. (2015). Increasing the adaptive capacity of organic farming systems in the face of climate change using action research methods. *Organic Agriculture*, pp. 1-13.

Buckingham, S., Anthony, S., Bellamy, P.H., Cardenas, L.M., Higgins, S., McGeough, K. & Topp, C.F.E. (2014). Review and analysis of global agricultural N2O emissions relevant to the UK. *Science of The Total Environment,* 487(0), pp. 164-172.

Bues, A., Preissel, S., Reckling, M., Zander, P., Kuhlmann, T., Topp, K., Watson, C., Lindström, K., Stoddard, F.L. & Murphy-Bokern, D. (2013). The Environmental Role of Protein Crops in the New Common Agricultural Policy. European Parliament, Directorate General for Internal Policies, Policy Department B: Structural and Cohesion Policies, Agricultural and Rural Development IP/B/AGRI/IC/2012-067. 112 pp. Access www.europarl.europa.eu/studies [29.10.2015].

Canfield, D.E., Glazer, A.N. & Falkowski, P.G. (2010). The Evolution and Future of Earth's Nitrogen Cycle. *Science,* 330(6001), pp. 192-196.

Carlsson, G. & Huss-Danell, K. (2003). Nitrogen fixation in perennial forage legumes in the field. *Plant and Soil,* 253(2), pp. 353-372.

Carof, M., Colomb, B. & Aveline, A. (2013). A guide for choosing the most appropriate method for multi-criteria assessment of agricultural systems according to decision-makers' expectations. *Agricultural Systems,* 115(0), pp. 51-62.

Castellazzi, M.S., Wood, G.A., Burgess, P.J., Morris, J., Conrad, K.F. & Perry, J.N. (2008). A systematic representation of crop rotations. *Agricultural Systems,* 97(1–2), pp. 26-33.

Catalogna, M., Dubois, M. & Navarrete, M. (2018). Diversity of experimentation by farmers engaged in agroecology. *Agronomy for Sustainable Development,* 38(5), p. 50.

Cernay, C., Ben-Ari, T., Pelzer, E., Meynard, J.-M. & Makowski, D. (2015). Estimating variability in grain legume yields across Europe and the Americas. *Scientific Reports,* 5, p. 11171.

Cernay, C., Makowski, D. & Pelzer, E. (2017). Preceding cultivation of grain legumes increases cereal yields under low nitrogen input conditions. *Environmental Chemistry Letters*.

Chalk, P.M. (1998). Dynamics of biologically fixed N in legume-cereal rotations: A review. *Australian Journal of Agricultural Research,* 49(3), pp. 303-316.

Clark, M. & Tilman, D. (2017). Comparative analysis of environmental impacts of agricultural production systems, agricultural input efficiency, and food choice. *Environmental Research Letters,* 12(6), p. 064016.

Cohen, J.E., Xu, M. & Schuster, W.S. (2012). Allometric scaling of population variance with mean body size is predicted from Taylor's law and density-mass allometry. *Proceedings of the National Academy of Sciences*, p. 201212883.

Cohen, J.E., Xu, M. & Schuster, W.S. (2013). Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. *Proceedings of the Royal Society B: Biological Sciences,* 280(1757), p. 20122955.

Colbach, N., Colas, F., Pointurier, O., Queyrel, W. & Villerd, J. (2017). A methodology for multi-objective cropping system design based on simulations. Application to weed management. *European Journal of Agronomy,* 87, pp. 59-73.

Crossa, J. (1988). A comparison of results obtained with two methods for assessing yield stability. *Theoretical and Applied Genetics,* 75(3), pp. 460-467.

De Notaris, C., Rasmussen, J., Sørensen, P. & Olesen, J.E. (2018). Nitrogen leaching: A crop rotation perspective on the effect of N surplus, field management and use of catch crops. *Agriculture, Ecosystems & Environment,* 255, pp. 1-11.

Debreczeni, K. & Körschens, M. (2003). Long-term field experiments of the world. *Archives of Agronomy and Soil Science,* 49(5), pp. 465-483.

Decourtye, A., Mader, E. & Desneux, N. (2010). Landscape enhancement of floral resources for honey bees in agro-ecosystems. *Apidologie,* 41(3), pp. 264-277.

Dehghani, H., Sabaghpour, S.H. & Sabaghnia, N. (2008). Genotype × environment interaction for grain yield of some lentil genotypes and relationship among univariate stability statistics. *Spanish Journal of Agricultural Research,* 6(3), pp. 385-394.

Descheemaeker, K., Ronner, E., Ollenburger, M., Franke, A.C., Klapwijk, C.J., Falconnier, G.N., Wichern, J. & Giller, K.E. (2016). Which options fit best? Operationalizing the socio-ecological niche concept. *Experimental Agriculture*, pp. 1-22.

DESTATIS (2017). Statistics of the Federal Statistical Office of Germany. "Landwirtschaftliche Bodennutzung, Anbau auf dem Ackerland, Fachserie 3 Reihe 3.1.2". Access https://www.destatis.de/ [03.01.2018].

Detlefsen, N.K. & Jensen, A.L. (2007). Modelling optimal crop sequences using network flows. *Agricultural Systems,* 94(2), pp. 566-572.

Dewhurst, R.J., Delaby, L., Moloney, A., Boland, T. & Lewis, E. (2009). Nutritive value of forage legumes used for grazing and silage. *Irish Journal of Agricultural and Food Research,* 48(2), pp. 167-187.

Di Matteo, J.A., Ferreyra, J.M., Cerrudo, A.A., Echarte, L. & Andrade, F.H. (2016). Yield potential and yield stability of Argentine maize hybrids over 45 years of breeding. *Field Crops Research,* 197, pp. 107-116.

Ditzler, L., Klerkx, L., Chan-Dentoni, J., Posthumus, H., Krupnik, T.J., Ridaura, S.L., Andersson, J.A., Baudron, F. & Groot, J.C.J. (2018). Affordances of agricultural systems analysis tools: A review and framework to enhance tool design and implementation. *Agricultural Systems,* 164, pp. 20-30.

Dogliotti, S., García, M.C., Peluffo, S., Dieste, J.P., Pedemonte, A.J., Bacigalupe, G.F., Scarlato, M., Alliaume, F., Alvarez, J., Chiappe, M. & Rossing, W.A.H. (2014). Co-innovation of family farm systems: A systems approach to sustainable agriculture. *Agricultural Systems,* 126, pp. 76-86.

Dogliotti, S., Rossing, W.A.H. & van Ittersum, M.K. (2003). rotat, a tool for systematically generating crop rotations. *European Journal of Agronomy,* 19(2), pp. 239-250.
Doré, T., Makowski, D., Malézieux, E., Munier-Jolain, N., Tchamitchian, M. & Tittonell, P. (2011). Facing up to the paradigm of ecological intensification in agronomy: Revisiting methods, concepts and knowledge. *European Journal of Agronomy,* 34(4), pp. 197-210.
Döring, T. & Reckling, M. *The adjusted coefficient of variation: An example calculation with global crop yield data for wheat*. Available at: DOI: 10.13140/RG.2.2.29086.43842.
Döring, T.F. (2015). Grain Legume Cropping Systems in Temperate Climates. In: De Ron, A.M. (ed. *Grain Legumes. Handbook of Plant Breeding*10). New York: Springer Science+Business Media, pp. 401-434.
Döring, T.F., Knapp, S. & Cohen, J.E. (2015). Taylor's power law and the stability of crop yields. *Field Crops Research,* 183, pp. 294-302.
Dusenbury, M.P., Engel, R.E., Miller, P.R., Lemke, R.L. & Wallander, R. (2008). Nitrous Oxide Emissions from a Northern Great Plains Soil as Influenced by Nitrogen Management and Cropping Systems. *Journal of Environmental Quality,* 37(2), pp. 542-550.
Eberhart, S.A. & Russell, W.A. (1966). Stability parameters for comparing varieties. *Crop Science,* 6(January-February), pp. 37-40.
Eckersten, H. (2017). *Cropping system research - A framework based on a literature review*. Uppsala, Sweden: Swedish University of Agricultural Sciences (SLU).
Eghball, B. & Power, J.F. (1995). Fractal description of temporal yield variability of 10 crops in the United States. *Agronomy Journal,* 87, pp. 152-156.
EIP-AGRI (2014). EIP-AGRI Focus Group Protein Crops: final report. Access http://ec.europa.eu/eip/agriculture/en/content/protein-crops [05.10.2018].
Eisenhauer, N., Milcu, A., Sabais, A.C.W., Bessler, H., Weigelt, A., Engels, C. & Scheu, S. (2009). Plant community impacts on the structure of earthworm communities depend on season and change with time. *Soil Biology and Biochemistry,* 41(12), pp. 2430-2443.
Eisler, Z., Bartos, I. & Kertész, J. (2008). Fluctuation scaling in complex systems: Taylor's law and beyond. *Advances in Physics,* 57(1), pp. 89-142.
Eriksen, J., Askegaard, M., Rasmussen, J. & Søegaard, K. (2015). Nitrate leaching and residual effect in dairy crop rotations with grass–clover leys as influenced by sward age, grazing, cutting and fertilizer regimes. *Agriculture, Ecosystems & Environment,* 212, pp. 75-84.
Ewert, F., van Ittersum, M.K., Heckelei, T., Therond, O., Bezlepkina, I. & Andersen, E. (2011). Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agriculture, Ecosystems & Environment,* 142(1), pp. 6-17.
Falconnier, G.N., Descheemaeker, K., Mourik, T.A.V. & Giller, K.E. (2016). Unravelling the causes of variability in crop yields and treatment responses for better tailoring of options for sustainable intensification in southern Mali. *Field Crops Research,* 187, pp. 113-126.
Falconnier, G.N., Descheemaeker, K., Van Mourik, T.A., Adam, M., Sogoba, B. & Giller, K.E. (2017). Co-learning cycles to support the design of innovative farm systems in southern Mali. *European Journal of Agronomy,* 89(Supplement C), pp. 61-74.
FAOSTAT (2014). Crop production data. In: Food and Agriculture Organization of the United Nations (ed. http://faostat3.fao.org/.

Farrington, J. & Martin, A.M. (1988). Farmer participatory research: A review of concepts and recent fieldwork. *Agricultural Administration and Extension,* 29(4), pp. 247-264.

Ferreira, D.F., Demétrio, C.G.B., Manly, B.F.J., de Almeida Machado, A. & Vencovsky, R. (2006). Statistical models in agriculture: biometrical methods for evaluating phenotypic stability in plant breeding. *Cerne,* 12(4), pp. 373-388.

Foyer, C.H., Lam, H.M., Nguyen, H.T., Siddique, K.H., Varshney, R.K., Colmer, T.D., Cowling, W., Bramley, H., Mori, T.A., Hodgson, J.M., Cooper, J.W., Miller, A.J., Kunert, K., Vorster, J., Cullis, C., Ozga, J.A., Wahlqvist, M.L., Liang, Y., Shou, H., Shi, K., Yu, J., Fodor, N., Kaiser, B.N., Wong, F.L., Valliyodan, B. & Considine, M.J. (2016). Neglecting legumes has compromised human health and sustainable food production. *Nat Plants,* 2, p. 16112.

Francis, T. & Kannenberg, L. (1978). Yield stability studies in short-season maize. I. A descriptive method for grouping genotypes. *Canadian Journal of Plant Science,* 58(4), pp. 1029-1034.

Franke, A.C., Baijukya, F., Kantengwa, S., Reckling, M., Vanlauwe, B. & Giller, K.E. (2016). Poor farmers – poor yields: Socio-economic, soil fertility and crop management indicators affecting climbing bean productivity in Northern Rwanda. *Experimental Agriculture,* FirstView, pp. 1-21.

Franke, A.C., van den Brand, G.J., Vanlauwe, B. & Giller, K.E. (2018). Sustainable intensification through rotations with grain legumes in Sub-Saharan Africa: A review. *Agriculture, Ecosystems & Environment,* 261, pp. 172-185.

Giller, K.E., Andersson, J.A., Corbeels, M., Kirkegaard, J., Mortensen, D., Erenstein, O. & Vanlauwe, B. (2015). Beyond conservation agriculture. *Frontiers in Plant Science,* 6(870).

Giller, K.E., Franke, A.C., Abaidoo, R., Baijukya, F.P., Bala, A., Boahen, S., Dashiell, K., Katengwa, S., Sanginga, J., Sanginga, N., Simmons, A., Turner, A., Woomer, P.L., Wolf, J.d. & Vanlauwe, B. (2014). N2Africa: Putting nitrogen fixation to work for smallholder farmers in Africa. In: Vanlauwe, B., van Asten, P. & Blomme, G. (eds) *Agro-ecological Intensification of Agricultural Systems in the African Highlands*. London: Routledge, pp. 156-174.

Giller, K.E., Tittonell, P., Rufino, M.C., van Wijk, M.T., Zingore, S., Mapfumo, P., Adjei-Nsiah, S., Herrero, M., Chikowo, R., Corbeels, M., Rowe, E.C., Baijukya, F., Mwijage, A., Smith, J., Yeboah, E., van der Burg, W.J., Sanogo, O.M., Misiko, M., de Ridder, N., Karanja, S., Kaizzi, C., K'Ungu, J., Mwale, M., Nwaga, D., Pacini, C. & Vanlauwe, B. (2011). Communicating complexity: Integrated assessment of trade-offs concerning soil fertility management within African farming systems to support innovation and development. *Agric. Syst.,* 104(2), pp. 191-203.

Goglio, P., Smith, W.N., Grant, B.B., Desjardins, R.L., Gao, X., Hanis, K., Tenuta, M., Campbell, C.A., McConkey, B.G., Nemecek, T., Burgess, P.J. & Williams, A.G. (2018). A comparison of methods to quantify greenhouse gas emissions of cropping systems in LCA. *Journal of Cleaner Production,* 172, pp. 4010-4017.

Graf, D.R.H., Zhao, M., Jones, C.M. & Hallin, S. (2016). Soil type overrides plant effect on genetic and enzymatic N2O production potential in arable soils. *Soil Biology and Biochemistry,* 100, pp. 125-128.

Groot, J.C.J., Oomen, G.J.M. & Rossing, W.A.H. (2012). Multi-objective optimization and design of farming systems. *Agricultural Systems,* 110, pp. 63-77.

Håkansson, S. (2003). *Weeds and weed management on arable land: an ecological approach.* Wallingford, UK: CABI Publishing.

Hamidov, A., Helming, K. & Balla, D. (2016). Impact of agricultural land use in Central Asia: a review. *Agronomy for Sustainable Development,* 36(1), p. 6.

Hawtin, G.C. & Hebblethwaite, P.D. (1983). Background and history of faba bean production. Chapter 1. In. (The Faba Bean (Vicia faba L.), a Basis for Improvement. London, UK: Butterworths, pp. 3-22.

Héault, C., Grossel, A., Mary, B., Roussel, M. & LÉOnard, J. (2012). Nitrous Oxide Emission by Agricultural Soils: A Review of Spatial and Temporal Variability for Mitigation. *Pedosphere,* 22(4), pp. 426-433.

Hernanz, J.L., Sanchez-Giron, V. & Navarrete, L. (2009). Soil carbon sequestration and stratification in a cereal/leguminous crop rotation with three tillage systems in semiarid conditions. *Agriculture, Ecosystems and Environment,* 133, pp. 114-122.

Hinton, N.J., Cloy, J.M., Bell, M.J., Chadwick, D.R., Topp, C.F.E. & Rees, R.M. (2015). Managing fertiliser nitrogen to reduce nitrous oxide emissions and emission intensities from a cultivated Cambisol in Scotland. *Geoderma Regional,* 4, pp. 55-65.

Hochman, Z., Gobbett, D.L. & Horan, H. (2017). Climate trends account for stalled wheat yields in Australia since 1990. *Global Change Biology,* 23(5), pp. 2071-2081.

Hoffmann, M.P., Haakana, M., Asseng, S., Höhn, J.G., Palosuo, T., Ruiz-Ramos, M., Fronzek, S., Ewert, F., Gaiser, T., Kassie, B.T., Paff, K., Rezaei, E.E., Rodríguez, A., Semenov, M., Srivastava, A.K., Stratonovitch, P., Tao, F., Chen, Y. & Rötter, R.P. (2018). How does inter-annual variability of attainable yield affect the magnitude of yield gaps for wheat and maize? An analysis at ten sites. *Agricultural Systems,* 159, pp. 199-208.

Huehn, M. (1990). Nonparametric measures of phenotypic stability. Part 1: Theory. *Euphytica,* 47(1), pp. 89-194.

Iannetta, P.P.M., Young, M., Bachinger, J., Bergkvist, G., Doltra, J., Lopez-Bellido, R.J., Monti, M., Pappa, V.A., Reckling, M., Topp, C.F.E., Walker, R.L., Rees, R.M., Watson, C.A., James, E.K., Squire, G.R. & Begg, G.S. (2016). A Comparative Nitrogen Balance and Productivity Analysis of Legume and Non-legume Supported Cropping Systems: The Potential Role of Biological Nitrogen Fixation. *Frontiers in Plant Science,* 7(1700).

IPCC (2006). 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston, H. S., Buendia, L., Miwa, K., Ngara, T., and Tanabe, K. (eds), Institute for Global Environmental Strategies (IGES), Japan.

Jensen, E.S., Peoples, M.B., Boddey, R.M., Gresshoff, P.M., Hauggaard-Nielsen, H., Alves, B.J.R. & Morrison, M.J. (2012). Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. *Agronomy for Sustainable Development,* 32(2), pp. 329-364.

Jensen, E.S., Peoples, M.B. & Hauggaard-Nielsen, H. (2010). Faba bean in cropping systems. *Field Crops Research,* 115(3), pp. 203-216.

Jeuffroy, M.H., Baranger, E., Carrouée, B., Chezelles, E.D., Gosme, M. & Hénault, C. (2013). Nitrous oxide emissions from crop rotations including wheat, oilseed rape and dry peas. *Biogeosciences,* 10.

Johnston, A.E. & Poulton, P.R. (2018). The importance of long-term experiments in agriculture: their management to ensure continued crop production and soil fertility; the Rothamsted experience. *European Journal of Soil Science,* 69(1), pp. 113-125.

Johnston, A.E., Poulton, P.R. & Coleman, K. (2009). Soil organic matter: its importance in sustainable agriculture and carbon dioxide fluxes. *Advances in agronomy,* 101, pp. 1-57.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Godfray, H.C.J., Herrero, M., Howitt, R.E., Janssen, S., Keating, B.A., Munoz-Carpena, R., Porter, C.H., Rosenzweig, C. & Wheeler, T.R. (2017). Brief history of agricultural systems modeling. *Agricultural Systems,* 155, pp. 240-254.

Kalkuhl, M., von Braun, J. & Torero, M. (2016). Volatile and extreme food prices, food security, and policy: an overview. Food Price Volatility and Its Implications for Food Security and Policy. In: Kalkuhl, M., von Braun, J. & Torero, M. (eds) *Food Price Volatility and Its Implications for Food Security and Policy* Springer Open, pp. 3-31.

Kermah, M., Franke, A.C., Adjei-Nsiah, S., Ahiabor, B.D.K., Abaidoo, R.C. & Giller, K.E. (2018). N2-fixation and N contribution by grain legumes under different soil fertility status and cropping systems in the Guinea savanna of northern Ghana. *Agriculture, Ecosystems & Environment,* 261, pp. 201-210.

Knapp, S. & van der Heijden, M.G.A. (2018). A global meta-analysis of yield stability in organic and conservation agriculture. *Nature Communications,* 9(1), p. 3632.

Knudsen, M.T., Hermansen, J.E., Olesen, J.E., Topp, C.F.E., Schelde, K., Angelopoulos, N. & Reckling, M. (2014). Climate impact of producing more grain legumes in Europe. *9th International Conference LCA of Food San Francisco, USA 8-10 October 2014*.

Kollas, C., Kersebaum, K.C., Nendel, C., Manevski, K., Müller, C., Palosuo, T., Armas-Herrera, C.M., Beaudoin, N., Bindi, M., Charfeddine, M., Conradt, T., Constantin, J., Eitzinger, J., Ewert, F., Ferrise, R., Gaiser, T., Cortazar-Atauri, I.G.d., Giglio, L., Hlavinka, P., Hoffmann, H., Hoffmann, M.P., Launay, M., Manderscheid, R., Mary, B., Mirschel, W., Moriondo, M., Olesen, J.E., Öztürk, I., Pacholski, A., Ripoche-Wachter, D., Roggero, P.P., Roncossek, S., Rötter, R.P., Ruget, F., Sharif, B., Trnka, M., Ventrella, D., Waha, K., Wegehenkel, M., Weigel, H.-J. & Wu, L. (2015). Crop rotation modelling – A European model intercomparison. *European Journal of Agronomy,* 70, pp. 98-111.

Köpke, U. & Nemecek, T. (2010). Ecological services of faba bean. *Field Crops Research,* 115(3), pp. 217-233.

Kravchenko, A.N., Snapp, S.S. & Robertson, G.P. (2017). Field-scale experiments reveal persistent yield gaps in low-input and organic cropping systems. *Proc Natl Acad Sci U S A,* 114(5), pp. 926-931.

Küchenmeister, F., Küchenmeister, K., Wrage, N., Kayser, M. & Isselstein, J. (2012). Yield and yield stability in mixtures of productive grassland species: Does species number or functional group composition matter? *Grassland science,* 58(2), pp. 94-100.

Lemaire, G., Gastal, F., Franzluebbers, A. & Chabbi, A. (2015). Grassland-Cropping Rotations: An Avenue for Agricultural Diversification to Reconcile High Production with Environmental Quality. *Environ Manage,* 56(5), pp. 1065-77.

LMC International (2009). *Evaluation of measures applied under the Common Agricultural Policy to the protein crop sector.* New York, Oxford, Kuala Lumpur.

Lobell, D., B. & Field, C., B. (2007). Global scale climate–crop yield relationships and the impacts of recent warming. *Environmental Research Letters,* 2(1), p. 014002.

Lobell, D.B., Bänziger, M., Magorokosho, C. & Vivek, B. (2011a). Nonlinear heat effects on African maize as evidenced by historical yield trials. *Nature Climate Change,* 1, p. 42.

Lobell, D.B., Schlenker, W. & Costa-Roberts, J. (2011b). Climate Trends and Global Crop Production Since 1980. *Science,* 333(6042), pp. 616-620.

López-Bellido, L., López-Bellido, R.J., Castillo, J.E. & López-Bellido, F.J. (2004). Chickpea response to tillage and soil residual nitrogen in a continuous rotation with wheat: I. Biomass and seed yield. *Field Crops Research,* 88(2-3), pp. 191-200.

Lorenz, M., Furst, C. & Thiel, E. (2013). A methodological approach for deriving regional crop rotations as basis for the assessment of the impact of agricultural strategies using soil erosion as example. *Journal of Environmental Management,* 127, pp. S37-47.

Lötjönen, S. & Ollikainen, M. (2017). Does crop rotation with legumes provide an efficient means to reduce nutrient loads and GHG emissions? *Review of Agricultural, Food and Environmental Studies,* 98(4), pp. 283-312.

Luetke-Entrup, N., Schneider, M. & Braun, J. (2003). Bearbeitungsgänge sparen, aber mit Augenmaß. *Neue Landwirtschaft,* Sonderheft Konservierende Bodenbearbeitung, pp. 9-14.

Lüscher, A., Mueller-Harvey, I., Soussana, J.F., Rees, R.M. & Peyraud, J.L. (2014). Potential of legume-based grassland–livestock systems in Europe: a review. *Grass and Forage Science,* 69(2), pp. 206-228.

Magrini, M.-B., Anton, M., Chardigny, J.-M., Duc, G., Duru, M., Jeuffroy, M.-H., Meynard, J.-M., Micard, V. & Walrand, S. (2018). Pulses for Sustainability: Breaking Agriculture and Food Sectors Out of Lock-In. *Frontiers in Sustainable Food Systems,* 2(64).

Magrini, M.-B., Anton, M., Cholez, C., Corre-Hellou, G., Duc, G., Jeuffroy, M.-H., Meynard, J.-M., Pelzer, E., Voisin, A.-S. & Walrand, S. (2016). Why are grain-legumes rarely present in cropping systems despite their environmental and nutritional benefits? Analyzing lock-in in the French agrifood system. *Ecological Economics,* 126, pp. 152-162.

Martin, G., Moraine, M., Ryschawy, J., Magne, M.-A., Asai, M., Sarthou, J.-P., Duru, M. & Therond, O. (2016). Crop–livestock integration beyond the farm level: a review. *Agronomy for Sustainable Development,* 36(3), p. 53.

Melander, B., Jabran, K., De Notaris, C., Znova, L., Green, O. & Olesen, J.E. (2018). Inter-row hoeing for weed control in organic spring cereals—Influence of inter-row spacing and nitrogen rate. *European Journal of Agronomy,* 101, pp. 49-56.

Meynard, J.-M., Charrier, F., Fares, M.h., Le Bail, M., Magrini, M.-B., Charlier, A. & Messéan, A. (2018). Socio-technical lock-in hinders crop diversification in France. *Agronomy for Sustainable Development,* 38(5), p. 54.

Müller-Stöver, D., Hauggaard-Nielsen, H., Eriksen, J., Ambus, P. & Johansen, A. (2012). Microbial biomass, microbial diversity, soil carbon storage, and stability after incubation of soil from grass–clover pastures of different age. *Biology and Fertility of Soils,* 48(4), pp. 371-383.

Müller, C., Elliott, J., Pugh, T.A.M., Ruane, A.C., Ciais, P., Balkovic, J., Deryng, D., Folberth, C., Cesar Izaurralde, R., Jones, C.D., Khabarov, N., Lawrence, P., Liu, W., Reddy, A.D., Schmid, E. & Wang, X. (2018). Global patterns of crop yield stability under additional nutrient and water inputs. *PLoS ONE,* 13(6).

Murphy-Bokern, D., Peeters, A. & Westhoek, H.J. (2017). The role of legumes in bringing protein to the table. In: D, M.-B., FL, S. & C, W. (eds) *Legumes in Cropping Systems* CABI Publishing, pp. 18-36.

Najafi, E., Devineni, N., Khanbilvardi, R.M. & Kogan, F. (2018). Understanding the Changes in Global Crop Yields Through Changes in Climate and Technology. *Earth's Future,* 6(3), pp. 410-427.

Naudin, K., Husson, O., Scopel, E., Auzoux, S., Giner, S. & Giller, K.E. (2015). PRACT (Prototyping Rotation and Association with Cover crop and no Till) – a tool for designing conservation agriculture systems. *European Journal of Agronomy,* 69, pp. 21-31.

Nemecek, T., von Richthofen, J.-S., Dubois, G., Casta, P., Charles, R. & Pahl, H. (2008). Environmental impacts of introducing grain legumes into European crop rotations. *European Journal of Agronomy,* 28(3), pp. 380-393.

Olesen, J.E., Trnka, M., Kersebaum, K.C., Skjelvåg, A.O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J. & Micale, F. (2011). Impacts and adaptation of European crop production systems to climate change. *European Journal of Agronomy,* 34(2), pp. 96-112.

Papendiek, F., Tartiu, V.E., Morone, P., Venus, J. & Hönig, A. (2016). Assessing the economic profitability of fodder legume production for Green Biorefineries – A cost-benefit analysis to evaluate farmers profitability. *Journal of Cleaner Production,* 112, pp. 3643-3656.

Pappa, V.A., Rees, R.M., Walker, R.L., Baddeley, J.A. & Watson, C.A. (2011). Nitrous oxide emissions and nitrate leaching in an arable rotation resulting from the presence of an intercrop. *Agriculture, Ecosystems & Environment,* 141(1-2), pp. 153-161.

Parajuli, R., Kristensen, I.S., Knudsen, M.T., Mogensen, L., Corona, A., Birkved, M., Peña, N., Graversgaard, M. & Dalgaard, T. (2017). Environmental life cycle assessments of producing maize, grass-clover, ryegrass and winter wheat straw for biorefinery. *Journal of Cleaner Production,* 142, pp. 3859-3871.

Peltonen-Sainio, P., Jauhiainen, L., Trnka, M., Olesen, J.E., Calanca, P., Eckersten, H., Eitzinger, J., Gobin, A., Kersebaum, K.C., Kozyra, J., Kumar, S., Marta, A.D., Micale, F., Schaap, B., Seguin, B., Skjelvåg, A.O. & Orlandini, S. (2010). Coincidence of variation in yield and climate in Europe. *Agriculture, Ecosystems & Environment,* 139(4), pp. 483-489.

Peltonen-Sainio, P. & Niemi, J.K. (2012). Protein crop production at the northern margin of farming: to boost or not to boost. *Agricultural and Food Science,* 21(4), pp. 370-383.

Pelzer, E., Bourlet, C., Carlsson, G., Lopez-Bellido, R.J., Jensen, E.S. & Jeuffroy, M.-H. (2017). Design, assessment and feasibility of legume-based cropping systems in three European regions. *Crop and Pasture Science,* 68(11), pp. 902-914.

Peoples, M.B., Brockwell, J., Herridge, D.F., Rochester, I.J., Alves, B.J.R., Urquiaga, S., Boddey, R.M., Dakora, F.D., Bhattarai, S., Maskey, S.L., Sampet, C., Rerkasem, B., Khan, D.F., Hauggaard-Nielsen, H. & Jensen, E.S. (2009a). The contributions of nitrogen-fixing crop legumes to the productivity of agricultural systems. *Symbiosis,* 48(1-3), pp. 1-17.

Peoples, M.B., Hauggaard-Nielsen, H. & Jensen, E.S. (2009b). The potential environmental benefits and risks derived from legumes in rotations. In: *Nitrogen Fixation in Crop Production*. (Agronomy Monograph, American Society of Agronomy, Crop Science Society of America, Soil Science Society of America,, pp. 349-385.

Peoples, M.B., Swan, A.D., Goward, L., Kirkegaard, J.A., Hunt, J.R., Li, G.D., Schwenke, G.D., Herridge, D.F., Moodie, M., Wilhelm, N., Potter, T., Denton, M.D., Browne, C., Phillips, L.A. & Khan, D.F. (2017). Soil mineral nitrogen benefits derived from legumes and comparisons of the apparent recovery of legume or fertiliser nitrogen by wheat. *Soil Research,* 55(6), pp. 600-615.

Persson, T., Bergkvist, G. & Kätterer, T. (2008). Long-term effects of crop rotations with and without perennial leys on soil carbon stocks and grain yields of winter wheat. *Nutrient cycling in agroecosystems,* 81(2), pp. 193-202.

PGRO (2018). Research Strategy for Combinable Pulse Crops of the UK Processors and Growers Research Organisation (PGRO). Access: www.pgro.org/index.php/applied-research-development/pulse-crop-r-d-strategy [05.10.2018].

Phelan, P., Moloney, A.P., McGeough, E.J., Humphreys, J., Bertilsson, J., O'Riordan, E. & O'Kiely, P. (2015). Forage legumes for grazing and conserving in ruminant production systems. *Crit Rev Plant Sci,* 34.

Philibert, A., Loyce, C. & Makowski, D. (2012). Quantifying uncertainties in N_2O emission due to N fertilizer application in cultivated areas. *PLoS ONE,* 7(11), p. e50950.

Piepho, H.P. (1998). Methods for Comparing the Yield Stability of Cropping Systems. *Journal of Agronomy and Crop Science,* 180(4), pp. 193-213.

Piepho, H.P., Büchse, A. & Emrich, K. (2003). A Hitchhiker's Guide to Mixed Models for Randomized Experiments. *Journal of Agronomy and Crop Science,* 189(5), pp. 310-322.

Pilbeam, C.J., Hebblethwaite, P.D. & Yusuf, A.A. (1990). Irrigation effects on the development and yield of determinate and indeterminate forms of autumn-sown faba bean. *Journal of the Science of Food and Agriculture,* 53(4), pp. 443-454.

Plaza-Bonilla, D., Nogué-Serra, I., Raffaillac, D., Cantero-Martínez, C. & Justes, É. (2018). Carbon footprint of cropping systems with grain legumes and cover crops: A case-study in SW France. *Agricultural Systems,* 167, pp. 92-102.

Plaza-Bonilla, D., Nolot, J.-M., Raffaillac, D. & Justes, E. (2015). Cover crops mitigate nitrate leaching in cropping systems including grain legumes: Field evidence and model simulations. *Agriculture, Ecosystems & Environment,* 212, pp. 1-12.

Popp, M., Rudstrom, M. & Manning, P. (2005). Spatial Yield Risk Across Region, Crop and Aggregation Method. *Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie,* 53(2-3), pp. 103-115.

Preissel, S., Reckling, M., Bachinger, J. & Zander, P. (2017). Introducing Legumes into European Cropping Systems: Farm-level Economic Effects. In: D, M.-B., FL, S. & C, W. (eds) *Legumes in Cropping Systems* CABI Publishing, pp. 209-225.

Preissel, S., Reckling, M., Schläfke, N. & Zander, P. (2015). Magnitude and farm-economic value of grain legume pre-crop benefits in Europe: A review. *Field Crops Research,* 175(0), pp. 64-79.

Profeta, A. & Hamm, U. (2019). Do consumers prefer local animal products produced with local feed? Results from a Discrete-Choice experiment. *Food Quality and Preference,* 71, pp. 217-227.

Prost, L., Reau, R., Paravano, L., Cerf, M. & Jeuffroy, M.-H. (2018). Designing agricultural systems from invention to implementation: the contribution of agronomy. Lessons from a case study. *Agricultural Systems,* 164, pp. 122-132.

R (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Raseduzzaman, M. & Jensen, E.S. (2017). Does intercropping enhance yield stability in arable crop production? A meta-analysis. *European Journal of Agronomy,* 91, pp. 25-33.

Ray, D.K., Gerber, J.S., MacDonald, G.K. & West, P.C. (2015). Climate variation explains a third of global crop yield variability. *Nature Communications,* 6, p. 5989.

Rebetzke, G.J., Fischer, R.A., van Herwaarden, A.F., Bonnett, D.G., Chenu, K., Rattey, A.R. & Fettell, N.A. (2014). Plot size matters: interference from intergenotypic competition in plant phenotyping studies. *Functional Plant Biology,* 41(2), pp. 107-118.

Reckling, M., Döring, T., Stein-Bachinger, K., Bloch, R. & Bachinger, J. (2015). Yield stability of grain legumes in an organically managed monitoring experiment. *Aspects of Applied Biology,* 128, pp. 57-62.

Reckling, M., Döring, T.F., Bergkvist, G., Chmielewski, F.-M., Stoddard, F.L., Watson, C.A., Seddig, S. & Bachinger, J. (2018). Grain legume yield instability has increased over 60 years in long-term field experiments as measured by a scale-adjusted coefficient of variation. *Aspects of Applied Biology,* 138, pp. 15-20.

Rees, R.M., Augustin, J., Alberti, G., Ball, B.C., Boeckx, P., Cantarel, A., Castaldi, S., Chirinda, N., Chojnicki, B., Giebels, M., Gordon, H., Grosz, B., Horvath, L., Juszczak, R., Klemedtsson, Å.K., Klemedtsson, L., Medinets, S., Machon, A., Mapanda, F., Nyamangara, J., Olesen, J., Reay, D., Sanchez, L., Sanz Cobena, A., Smith, K.A., Sowerby, A., Sommer, M., Soussana, J.F., Stenberg, M., Topp, C.F.E., van Cleemput, O., Vallejo, A., Watson, C.A. & Wuta, M. (2012). Nitrous oxide emissions from European agriculture; an analysis of variability and drivers of emissions from field experiments. *Biogeosciences Discussions,* 9(7), pp. 9259-9288.

Robson, M.C., Fowler, S.M., Lampkin, N.H., Leifert, C., Leitch, M., Robinson, D., Watson, C.A. & Litterick, A.M. (2002). The agronomic and economic potential of break crops for ley/arable rotations in temperate organic agriculture. *Advances in Agronomy,* Volume 77, pp. 369-427.

Rochette, P. & Janzen, H.H. (2005). Towards a Revised Coefficient for Estimating N2O Emissions from Legumes. *Nutrient Cycling in Agroecosystems,* 73(2), pp. 171-179.

Ronner, E. (2018). *From targeting to tailoring: Baskets of options for legume cultivation among African smallholders.* Diss. Wageningen, the Netherlands: Wageningen University.

Ronner, E., Franke, A.C., Vanlauwe, B., Dianda, M., Edeh, E., Ukem, B., Bala, A., van Heerwaarden, J. & Giller, K.E. (2016). Understanding variability in soybean yield and response to P-fertilizer and rhizobium inoculants on farmers' fields in northern Nigeria. *Field Crops Research,* 186, pp. 133-145.

Röös, E., Carlsson, G., Ferawati, F., Hefni, M., Stephan, A., Tidåker, P. & Witthöft, C. (2018). Less meat, more legumes: prospects and challenges in the transition toward sustainable diets in Sweden. *Renewable Agriculture and Food Systems*, pp. 1-14.

Rötter, R.P., Appiah, M., Fichtler, E., Kersebaum, K.C., Trnka, M. & Hoffmann, M.P. (2018). Linking modelling and experimentation to better capture crop impacts of agroclimatic extremes—A review. *Field Crops Research,* 221, pp. 142-156.

Rubiales, D. & Fernández-Aparicio, M. (2012). Innovations in parasitic weeds management in legume crops. A review. *Agronomy for Sustainable Development,* 32(2), pp. 433-449.

Rubiales, D., Fondevilla, S., Chen, W., Gentzbittel, L., Higgins, T.J.V., Castillejo, M.A., Singh, K.B. & Rispail, N. (2015). Achievements and Challenges in Legume Breeding for Pest and Disease Resistance. *Critical Reviews in Plant Sciences,* 34(1-3), pp. 195-236.

Rudstrom, M., Popp, M., Manning, P. & Gbur, E. (2002). Data aggregation issues for crop yield risk analysis. *Canadian Journal of Agricultural Economics,* 50(2), pp. 185-200.

Sabais, A.C.W., Scheu, S. & Eisenhauer, N. (2011). Plant species richness drives the density and diversity of Collembola in temperate grassland. *Acta Oecologica,* 37(3), pp. 195-202.

Sadok, W., Angevin, F., Bergez, J.-É., Bockstaller, C., Colomb, B., Guichard, L., Reau, R. & Doré, T. (2009a). Ex ante Assessment of the Sustainability of Alternative Cropping Systems: Implications for Using Multi-criteria Decision-Aid Methods - A Review. In: Lichtfouse, E., Navarrete, M., Debaeke, P., Véronique, S. & Alberola, C. (eds) *Sustainable Agriculture* Springer Netherlands, pp. 753-767.

Sadok, W., Angevin, F., Bergez, J.-E., Bockstaller, C., Colomb, B., Guichard, L., Reau, R., Messéan, A. & Doré, T. (2009b). MASC, a qualitative multi-attribute decision model for ex ante assessment of the sustainability of cropping systems. *Agronomy for Sustainable Development,* 29(3), pp. 447-461.

Schmidt, P., Möhring, J., Koch, R.J. & Piepho, H.-P. (2018). More, Larger, Simpler: How Comparable Are On-Farm and On-Station Trials for Cultivar Evaluation? *Crop Science,* 58(4), pp. 1508-1518.

Schönhart, M., Schmid, E. & Schneider, U.A. (2011). CropRota – A crop rotation model to support integrated land use assessments. *European Journal of Agronomy,* 34(4), pp. 263-277.

Shen, J., Yuan, L., Zhang, J., Li, H., Bai, Z. & Chen, X. (2011). Phosphorus dynamics: from soil to plant. *Plant Physiol,* 156.

Sinclair, F.L. (2017). Systems science at the scale of impact: Reconciling bottom up participation with the production of widely applicable research outputs. In: Oborn, I., Vanlauwe, B., Phillips, M., Thomas, R., Brooijmans, W. & Atta-Krah, K. (eds) *Sustainable Intensification in Smallholder Agriculture*. London: Routledge.

St-Martin, A., Vico, G., Bergkvist, G. & Bommarco, R. (2017). Diverse cropping systems enhanced yield but did not improve yield stability in a 52-year long experiment. *Agriculture, Ecosystems & Environment,* 247(Supplement C), pp. 337-342.

Stagnari, F., Maggio, A., Galieni, A. & Pisante, M. (2017). Multiple benefits of legumes for agriculture sustainability: an overview. *Chemical and Biological Technologies in Agriculture,* 4(1), p. 2.

Stein-Bachinger, K. & Fuchs, S. (2012). Protection strategies for farmland birds in legume–grass leys as trade-offs between nature conservation and farmers' needs. *Organic Agriculture,* 2(2), pp. 145-162.

Stein, S. & Steinmann, H.-H. (2018). Identifying crop rotation practice by the typification of crop sequence patterns for arable farming systems – A case study from Central Europe. *European Journal of Agronomy,* 92, pp. 30-40.

Steinmann, H.-H. & Dobers, E. (2013). Spatio-temporal analysis of crop rotations and crop sequence patterns in Northern Germany: potential implications on plant health and crop protection. *Journal of Plant Diseases and Protection,* 120(2), pp. 85–94.

Stoddard, F.L., Balko, C., Erskine, W., Khan, H.R., Link, W. & Sarker, A. (2006). Screening techniques and sources of resistance to abiotic stresses in cool-season food legumes. *Euphytica,* 147(1), pp. 167-186.

Stützel, H., Brüggemann, N. & Inzé, D. (2016). The Future of Field Trials in Europe: Establishing a Network Beyond Boundaries. *Trends in Plant Science,* 21(2), pp. 92-95.

Sumberg, J., Okali, C. & Reece, D. (2003). Agricultural research in the face of diversity, local knowledge and the participation imperative: theoretical considerations. *Agricultural Systems,* 76(2), pp. 739-753.

Suter, M., Finn, J.A., Connolly, J., Loges, R. & Lüscher, A. (2015). Gain in Nitrogen Yield from Grass-Legume Mixtures is Robust Over a Wide Range of Legume Proportions and Environmental Conditions. *Procedia Environmental Sciences,* 29, pp. 187-188.

Taylor, L. (1961). Aggregation, variance and the mean. *Nature,* 189, pp. 732–735.

Tidåker, P., Sundberg, C., Öborn, I., Kätterer, T. & Bergkvist, G. (2014). Rotational grass/clover for biogas integrated with grain production – A life cycle perspective. *Agricultural Systems,* 129, pp. 133-141.

Tigchelaar, M., Battisti, D.S., Naylor, R.L. & Ray, D.K. (2018). Future warming increases probability of globally synchronized maize production shocks. *Proceedings of the National Academy of Sciences*.

Topp, F.E., Reckling, M., Hanegraaf, M., Bachinger, J., Walker, R.L., Buckingham, S., Sykes, A.J. & Watson, A. (2017). Farmer friendly tools – how can they help support decision making under a changing climate? *Aspects of Applied Biology,* 136, pp. 1-6.

Trnka, M., Brázdil, R., Olesen, J.E., Eitzinger, J., Zahradníček, P., Kocmánková, E., Dobrovolný, P., Štěpánek, P., Možný, M., Bartošová, L., Hlavinka, P., Semerádová, D., Valášek, H., Havlíček, M., Horáková, V., Fischer, M. & Žalud, Z. (2012). Could the changes in regional crop yields be a pointer of climatic change? *Agricultural and Forest Meteorology,* 166-167, pp. 62-71.

TRUE (2018). Meat substitutes and lentil pasta: Legume products on the rise in Europe. *Press Release, 09.02.2018.* TRansition paths to sUstainable legume based systems in Europe, p. 2.

Urruty, N., Tailliez-Lefebvre, D. & Huyghe, C. (2016). Stability, robustness, vulnerability and resilience of agricultural systems. A review. *Agronomy for Sustainable Development,* 36(1), p. 15.

van der Burgt, G.J.H.M., Oomen, G.J.M., Habets, A.S.J. & Rossing, W.A.H. (2006). The NDICEA model, a tool to improve nitrogen use efficiency in cropping systems. *Nutrient Cycling in Agroecosystems,* 74(3), pp. 275-294.

van Ittersum, M.K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., van der Wal, T., Wien, J.E. & Wolf, J. (2008). Integrated assessment of agricultural systems – A component-based framework for the European Union (SEAMLESS). *Agricultural Systems,* 96(1), pp. 150-165.

van Vugt, D., Franke, A.C. & Giller, K.E. (2018). Understanding variability in the benefits of N2-fixation in soybean-maize rotations on smallholder farmers' fields in Malawi. *Agriculture, Ecosystems & Environment,* 261, pp. 241-250.

Voisin, A.-S., Guéguen, J., Huyghe, C., Jeuffroy, M.-H., Magrini, M.-B., Meynard, J.-M., Mougel, C., Pellerin, S. & Pelzer, E. (2014). Legumes for feed, food, biomaterials and bioenergy in Europe: a review. *Agronomy for Sustainable Development,* 34(2), pp. 361-380.

Voisin, A.-S., Salon, C., Munier-Jolain, N.G. & Ney, B. (2002). Quantitative effects of soil nitrate, growth potential and phenology on symbiotic nitrogen fixation of pea (Pisum sativum L.). *Plant and Soil,* 243(1), pp. 31-42.

Von Richthofen, J.-S., Pahl, H., Bouttet, D., Casta, P., Cartrysse, C., Charles, R. & Lafarga, A. (2006a). What do European farmers think about grain legumes. *Grain Legumes,* 45, pp. 14-15.

Von Richthofen, J.-S., Pahl, H. & Nemecek, T. (2006b). *Economic interest of grain legumes in European crop rotations*: GL-Pro.

Watson, C., Reckling, M., Preissel, S., Bachinger, J., Bergkvist, G., Kuhlman, T., Lindström, K., Nemecek, T., Topp, C., Vanhatalo, A., Zander, Z., Murphy-Bokern, D. & Stoddard, F. (2017). Grain legume production and use in European agricultural systems. *Advances in Agronomy,* 144(1), pp. 235-303.

Webber, H., Ewert, F., Olesen, J.E., Müller, C., Fronzek, S., Ruane, A.C., Bourgault, M., Martre, P., Ababaei, B., Bindi, M., Ferrise, R., Finger, R., Fodor, N., Gabaldón-Leal, C., Gaiser, T., Jabloun, M., Kersebaum, K.-C., Lizaso, J.I., Lorite, I.J., Manceau, L., Moriondo, M., Nendel, C., Rodríguez, A., Ruiz-Ramos, M., Semenov, M.A., Siebert, S., Stella, T., Stratonovitch, P., Trombi, G. & Wallach, D. (2018). Diverging importance of drought stress for maize and winter wheat in Europe. *Nature Communications,* 9(1), p. 4249.

Westphal, C., Steffan-Dewenter, I. & Tscharntke, T. (2003). Mass flowering crops enhance pollinator densities at a landscape scale. *Ecology Letters,* 6(11), pp. 961-965.

Zander, P., Amjath-Babu, T.S., Preissel, S., Reckling, M., Bues, A., Schläfke, N., Kuhlman, T., Bachinger, J., Uthes, S., Stoddard, F., Murphy-Bokern, D. & Watson, C. (2016). Grain legume decline and potential recovery in European agriculture: a review. *Agronomy for Sustainable Development,* 36(2), p. 26.

Zimmer, S., Liebe, U., Didier, J.-P. & Heß, J. (2016). Luxembourgish farmers' lack of information about grain legume cultivation. *Agronomy for Sustainable Development,* 36(2), p. 2.

Acknowledgements

Göran, I will not forget our conversation at one *Legume-Futures* meeting about a potential PhD project and your offer to supervise me. I thank you for that! Although it was quite difficult 'finding money' for the 4-year period, you always made me look positive into the future. You were always supporting and gave me the guidance that I needed and the freedom to develop my ideas.

Hans, I learned from you how farmers grow legumes in Europe and how to think in a crop rotation perspective (and to put this thinking into models). You gave the right practical input to the agronomic questions and I thank you for the space to work independently and to support me all the time.

Christine & Fred, I always felt you being near supporting my PhD although you were hundreds of kilometres away. I thank you for the help developing the ideas into papers and for not getting tired teaching me academic English.

Thomas, you trained me how to assess yield stability with novel methods and to think critically about everything being 'felt' related to this topic that has not been empirically tested. I thank you for the great teamwork.

The contribution from all **co-authors** to the papers in this thesis is much appreciated, especially from **Martin** for programming the rotation generator, **Kairsty** for the support with calculating the N_2O emissions, **Robin**, **Ion** and **Aurelio** for you active contribution to the cropping system assessment.

At ZALF, I want to thank **Sonoko**, **Peter**, **Karin** and **Angelika** for your support that made this thesis possible. I want to thank **Ralf** for your innovative ideas, discussions about involving farmers in our research and on how to deal with the work-life-balance. **Guni**, I want to thank you for the practical support and the very careful data processing and analysis. I thank **Gerlinde**, and the staff from the research station, especially **Sven**, **Dietmar** and **Brigitte**, for managing the trials with often very special wishes and **Dr. Roth** and **Johannes** for the discussions around agronomy. **Sigi**, I thank you for the maps. There are many more people that supported me at different times and I like to thank them for that.

At SLU, I would like to thank **Pierre**, **Marcos**, **Elsa**, **Nina** & **Raj**, **Brigitta**, **Bodil**, **Paula**, **Kristin**, **Thomas**, **Janne** and **Johannes** for making me feel part of *Institutionen för växtproduktionsekologi* and always welcome in Sweden. **Ingrid**, I thank you for supporting me during the first part of the thesis.

This work would not have been possible without the contribution from many **farmers** and **advisors**, as well as the **funding** through *Legume Futures* and *Climate-CAFE*.