THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Agricultural nutrient budgets in Europe: data, methods, and indicators

RASMUS EINARSSON

Department of Space, Earth and Environment
Division of Physical Resource Theory
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2020

Agricultural nutrient budgets in Europe: data, methods, and indicators
RASMUS EINARSSON

ISBN 978-91-7905-367-3

Doktorsavhandlingar vid Chalmers tekniska högskola, Ny serie nr. 4834
ISSN 0346-718X

Department of Space, Earth and Environment
Division of Physical Resource Theory
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-772 1000

Typeset using LaTeX.

Printed by Chalmers Reproservice
Göteborg, Sweden 2020

Abstract

Agricultural production systems feed humanity but also cause a range of adverse environmental effects, including climate change, loss of biodiversity, and pollution of air and water. A main cause of these effects is the emissions of nitrogen (N) and phosphorus (P) that occur as a side effect of nutrient cycling in agriculture. One of the things that is needed to mitigate N and P pollution is a quantitative understanding of N and P flows in agricultural systems. A common tool for this is the nutrient budget. A nutrient budget quantifies inputs and outputs of nutrients in a system and can be used to understand how the system functions as well as to calculate quantitative environmental indicators for farms, regions, or products.

This thesis aims to explore and expand the limits of how agricultural N and P budgets can be used to support environmental research and decision-making, focusing on European agriculture. To this end, the thesis looks into two broad research questions: (1) What are the limits to the accuracy and level of detail that can be attained in N and P budgets of European agricultural systems? (2) How are present and proposed uses of agricultural N and P budgets and derived indicators limited by (a) the inherent property that agricultural nutrient budgets do not account for environmental impacts, and (b) by uncertainties and lack of data in the estimation of nutrient budgets?

This thesis builds on five appended research papers that explore various aspects of data sources, uncertainties, and possible uses of N and P budgets in Europe. International and national data sources are scrutinized and used to estimate N and P budgets. Novel ways to combine existing data sources are explored. The use of nutrient budgets with various system boundaries, with different degrees of spatial resolution, and in different time periods is discussed, emphasizing that the best approach is not only a question of data supply but also of intended audience and purpose.

Keywords: nitrogen, phosphorus, nutrient budget, nutrient balance, agriculture, environment, indicators, Europe, EU

Appended publications

This thesis is based upon the following papers:

Paper A R. Einarsson and U. M. Persson (2017). Analyzing key constraints to biogas production from crop residues and manure in the EU—A spatially explicit model. *PLOS ONE* **12** (1), e0171001. DOI: `10.1371/journal.pone.0171001`.

Paper B R. Einarsson, C. Cederberg, and J. Kallus (2018). Nitrogen flows on organic and conventional dairy farms: a comparison of three indicators. *Nutrient Cycling in Agroecosystems* **110** (1), pp. 25–38. DOI: `10.1007/s10705-017-9861-y`.

Paper C R. Einarsson and C. Cederberg (2019). Is the nitrogen footprint fit for purpose? An assessment of models and proposed uses. *Journal of Environmental Management* **240**, pp. 198–208. DOI: `10.1016/j.jenvman.2019.03.083`.

Paper D R. Einarsson, D. Pitulia, and C. Cederberg (2020). Subnational nutrient budgets to monitor environmental risks in EU agriculture: calculating phosphorus budgets for 243 EU28 regions using public data. *Nutrient Cycling in Agroecosystems.* DOI: `10.1007/s10705-020-10064-y`.

Paper E R. Einarsson, A. Sanz Cobeña, E. Aguilera, G. Billen, H. J. M. van Grinsven, and L. Lassaletta (2020). "Crop production and nitrogen use in European cropland and grassland 1961-2013". *To be submitted for publication.*

Author contributions

Paper A: RE conceived the idea. RE collected data, designed the model, and made the calculations. RE and UMP interpreted the results and wrote the paper.

Paper B: CC and RE conceived the idea. RE selected and cleaned data and made the calculations with input from CC. RE and JK designed the statistical analysis. RE wrote the paper with input from CC and JK.

Paper C: RE and CC conceived the idea. RE collected and analyzed data about N footprint models. RE and CC interpreted results. RE wrote the paper with input from CC.

Paper D: CC and RE conceived the idea. DP and RE contributed equally to data collection and analysis. CC, DP, and RE interpreted results. RE implemented the final version of the model. RE and DP wrote the paper with input from CC.

Paper E: RE and LL conceived the idea. RE collected and cleaned data in close collaboration with LL, ASC, and EA. RE, LL, ASC, and EA designed the model with inputs from GB, JG, and HvG. RE drafted the manuscript with input from the others.

Acknowledgements

I want to thank everyone who made this experience so enjoyable. My supervisor and co-author Christel Cederberg for guidance and support, and for encouraging me from the very beginning to pursue my own research ideas. Kristian Lindgren for urging me to apply to the doctoral program in the first place and for many memorable moments together. Martin Persson who also talked me into applying and guided me through the process with the first publication. My co-supervisor Daniel Johansson for guidance and encouragement. I still want to build that statistical model we talked about, and would love to have your input even if you are now officially off the hook. Daniel Pitulia for your friendly and constructive attitude working together. Jonatan Kallus for companionship through the ups and downs of doctoral studies, and for many wonderful conversations where you understood my point before I even finished the sentence.

Göteborgs Handelskompanis deposition for funding this research and the people at Chalmers who ensured the freedom that Laura and I have enjoyed.

Co-authors and friends in Spain: Luis Lassaletta for responding to my unsolicited and inquisitive questions with excitement and an invitation to Madrid. Alberto Sanz Cobeña and Eduardo Aguilera for excellent work meetings and unending patience with all the details. Ivanka Puigdueta Bartolomé for friendship and guidance in everyday life at CEIGRAM. Thank you, all four, for inviting me into your lives and homes.

Gilles Billen, Josette Garnier, Hans van Grinsven, Maria Henriksson, Markus Hoffman, Michelle McCrackin, Mark Sutton, and Annika Svanbäck for enjoyable and stimulating conversations. Everyone at the Focus on Nutrients project who have supported our research. All the land use researchers at Chalmers for creating such a nice learning environment.

Sara Jutterström, Filip Moldan, and Magdalena Wallman for showing a glimpse of the reality of environmental monitoring and experimental data collection. It made me feel more gratitude for all the data in the world.

Everyone at Physical Resource Theory for the warm and lively atmosphere. My long-time office mate Stefan Åström for the abrupt mix of silence and conversation. Friends and family for everything you do, and especially Esmeralda for patience and encouragement along the way.

Contents

CHAPTER 1

Introduction

Agriculture has altered the Earth's environment for millennia. What is new is the scale and intensity. Agriculture today feeds almost eight billion people but, ironically, its adverse environmental effects are increasingly identified as a major threat to long-term human prosperity [1, 2].

A main driver of agriculture's increased productivity as well as its adverse environmental effects has been the increased supply of plant nutrients in the form of synthetic nitrogen (N) and mined phosphorus (P) fertilizers. Starting from almost zero in year 1900, the global inputs of fertilizer N and P have grown to almost equal the outputs recovered in plant production and far exceed the amounts in human food intake [3–5]. The fertilizers are used in crop production but have reshaped the whole agricultural system by increasing the flows of livestock feed, manure, and crop residues to magnitudes that hardly could be imagined just fifty or a hundred years ago. The unwanted side effect is that large quantities of N and P escape into the environment, contributing to an array of environmental problems including climate change, biodiversity loss, air pollution, harmful algal blooms and hypoxic waters, stratospheric ozone depletion, and toxic groundwater contamination [6–10].

There are no simple solutions in sight for these problems. The issue of nutrient cycling is only one in a web of interconnected issues around agriculture and food: providing a growing world population with nutritious and culturally acceptable food; maintaining and improving soil quality; conserving forests and other threatened ecosystems; securing the livelihoods of the world's farmers; mitigating and adapting to climate change; managing limited freshwater resources; and so on. There is broad scientific and political agreement that change is needed in our agriculture and food systems [11, 12], but no obvious answer to what principles, priorities, and methods should be used to guide that change.

Even within the narrower issue of nutrient cycling, it is possible to emphasize quite different approaches. Improved technology and management can improve the efficiency of nutrient use and thereby reduce both inputs and environmental pollution. Shifting human diets to more plant-based food would reduce the need for livestock production, which consumes the majority of the world's crop production, yet supplies a minority of global food calories and proteins [9, 13]. Prioritizing mitigation in sensitive environments, maybe even suspending or limiting agriculture in some areas, may seem unfair but would be the most cost efficient approach given the heterogeneity of both land use and environmental impacts [14, 15]. In principle, all these approaches have support in research and policy, but priorities diverge and most likely will continue to do so.

In any case, sound decisions will require a good quantitative understanding of N and P flows in agricultural systems. A common tool for this is the nutrient budget. A nutrient budget quantifies inputs and outputs of nutrients in a system. In a crop field, for example, the inputs are things like fertilizers, manure, and atmospheric deposition of N or P, and the output is the N or P removed in the harvest. The difference between the combined inputs and outputs is called balance or surplus, and roughly corresponds to the N or P lost to the environment, adjusted for changes in soil nutrient storage [16]. The nutrient surplus is commonly expressed as an average flow per unit area and time (e.g. $kg\,ha^{-1}\,y^{-1}$) and thus acts as an indicator of the potential pollution intensity. Another common indicator is the nutrient use efficiency (NUE) = outputs / inputs, a measure of how efficiently a system transforms inputs into useful outputs.

The nutrient budget and these derived indicators are used for several purposes. In farm management, nutrient budgets on farm or field level can be used for planning and for benchmarking agronomic and environmental performance [17, 18]; in official OECD statistics and in evaluation of EU policies, national N and P balances are used as agri-environmental indicators [19–21]; and in agricultural and environmental research, nutrient budgets in many variations are used to analyze farms, fields, production systems, watersheds, countries, and so on [22–25].

As the concerns over agriculture's environmental impacts have intensified, so has the research on nutrient budgets and derived indicators as tools for research and decision-making. Methods have been scrutinized and clarified, national and international datasets have grown, and there have been efforts to assess and improve the environmental relevance of budgets and derived indicators. This thesis identifies and connects to four interrelated strands of research in this area.

One strand of research has addressed the issue that national nutrient budgets by definition do not show any subnational heterogeneity and thus may fail to identify local environmental risks. This happens, for example, when one region is specialized in highly efficient crop production and another in intensive livestock production. The livestock region may have a large excess supply of manure, causing locally severe environmental effects, but when averaged with the crop region the combined balance might cause no concern [26]. Therefore, researchers have developed methods and datasets to increase the geographical resolution of N and P budgets, confirming the need for continued work on subnational nutrient budgets [27–31].

Another strand of research has shifted the attention from geographical system boundaries, such as fields, farms, and regions, to instead create nutrient budgets and indicators for products or production systems. For example, livestock farms that purchase livestock feed thereby externalize the environmental costs associated with cultivation of the feed crops [22]. Researchers have therefore explored how to draw an alternative system boundary around the system behind a product, such as a livestock production system including the feed production and other related processes [23, 32–34]. While this approach leads to a range of new technical complications, it deserves continued development as it provides new information about the environmental risks associated with different products.

A third strand of research estimates and studies historical nutrient budgets for agricultural and food systems. One example is a global time series of country-level N budgets reaching back to 1961 [35], and another covers 33 subnational regions of France back to 1852 [25]. These studies do not only deepen our understanding of how agricultural and food systems have changed in the past, but arguably can also be used to understand the options for future change [4, 36, 37].

The fourth strand of research looks into how nutrient balances and other budget-based indicators relate to environmental effects. This relation is both variable and uncertain since N and P are transported, transformed, and stored in the environment in ways that depend on soil properties, topography, climate, weather events, and other factors, but which are only partly understood [38, 39]. For both N and P, this raises questions about what the nutrient balance or NUE should be used for, but it does not invalidate them altogether: There *are* links between the indicators and environmental effects, but the precise *quality* of those links is complicated [16, 40–42]. The appropriate uses of the nutrient balance, NUE, and related indicators thus remains a question of scientific inquiry as well as a question of intended purpose.

These examples illustrate that nutrient budget research can conceptually be divided into two parts, one about the agricultural systems and the other about the budgets and indicators as such. The first part uses budgets and indicators as a toolbox to describe and learn about agricultural systems: How do production systems compare in terms of nutrient budgets? How have budgets changed over time? What can be done to decrease surpluses or increase NUE? The second part takes a step back to inspect and improve the tools: How can budgets be accurately estimated? What questions can they answer, depending on system boundaries, geographical resolution, and so on? What can the various agri-environmental indicators meaningfully be used for? Much of the research, including this thesis, does a bit of both. On one hand, it is important to think critically about the tools when using them, and on the other, it helps to use the tools to recognize and address their shortcomings.

The main focus of this thesis, however, is on the second part, the study of agricultural nutrient budgets and derived indicators as a toolbox under development. In the pages that follow, I describe and reflect upon the work my co-authors and I have done to improve these tools and the ways they are used. It is my hope that the data, methods, and reflections I present will be useful for those who use nutrient budgets and related indicators as a basis for decisions about the future of our food systems.

1.1 Aim and research questions

This thesis aims to explore and expand the limits of how European agricultural N and P budgets can be used to support environmental research and decision-making.

The thesis builds on five appended research papers that all contribute to this aim in different ways. While the papers address their own and loosely related research questions, this thesis is an attempt to synthesize our findings and to address the following research questions:

RQ1 What are the limits to the accuracy and level of detail that can be attained in N and P budgets of European agricultural systems? Specifically:

  a. What geographical resolution can be attained in budgets of subnational regions?
  b. How can budgets be estimated with various alternative system boundaries proposed in the literature?
  c. What time period can be covered in subnational and national budgets?

RQ2 How are present and proposed uses of agricultural N and P budgets limited by

  a. the inherent limitation that budgets do not account for the fate of N and P beyond the agricultural system?
  b. the less fundamental limitations brought about by uncertainties and lack of data?

Most of the findings in this thesis are limited to the study of European agricultural systems, more specifically the territory of the former EU28. The main reason for this is that several important data sources cover only EU countries. Another reason is that Europe has a fairly homogeneous climate and agricultural structure compared to the variation of the rest of the world. This does not mean that the findings in this thesis are useless outside Europe, but considering the specificity of many datasets and assumptions, it is appropriate to point out that a majority of the conclusions are specific to Europe.

1.2 Structure of the thesis

This introduction chapter ends in Section 1.3 with an overview of the five appended papers and how they relate to the aim and research questions of the thesis. Then, Chapter 2 presents some background on the N and P cycles and on nutrient budgets and derived indicators. Chapter 2 is intended as an introduction for someone who is not already familiar with the research field, perhaps a researcher or practitioner in a related field, or a PhD student at the start of their career. Chapter 3 is about the research approach I have used, focusing on some methodological issues rather than specific method choices and descriptions, which are already documented in the appended papers. Chapter 4 presents the main findings and some reflections on our work. Chapter 5 concludes the thesis with a few final reflections.

1.3 Summary of appended papers

The five appended papers are the main research outputs I have contributed to in my time as a PhD student. Here I give an overview what they are about and how they relate to the aim and research questions of this thesis. Table 1.1 gives a condensed overview of the appended papers and their main contributions to the research questions of the thesis.

1.3.1 Paper A

Aim and approach

This paper is about the potential for biogas production from crop residues and manure in the EU. To this end, we constructed a geographically explicit model of these substrates based on a range of international datasets: Eurostat's subnational datasets on crop and livestock production, FAO's Gridded Livestock of the World v2 dataset, EEA's Corine Land Cover dataset, and data

Table 1.1: Overview the appended papers and their relation to the research questions of this thesis (see page 4).

	Paper A	Paper B	Paper C	Paper D	Paper E
Paper topic	Potential for biogas production from crop residues and manure in the EU.	N surplus and footprint indicators for farms vs. production systems.	Critical assessment of N footprint models and use cases.	Subnational vs. national P budgets in EU28.	Country-level N budgets in European cropland 1961–2013.
RQ 1a	Estimation of N flows in crop residues and manure with resolution ~1 km.			Estimation of P budgets for 243 subnational regions in the EU.	
RQ 1b		Estimation of farm and production system budgets.		Model variant excluding extensively managed grassland.	Data fusion and modeling to exclude permanent grassland.
RQ 1c				Explicit aim to use regularly updated international datasets.	Data fusion of FAOSTAT, Eurostat, and national data sources back to 1961.
RQ 2a			Analysis of the meaning, use cases, and “fitness for purpose” of N footprints.		
RQ 2b		Uncertainty analysis. Discussion on appropriate uses of the indicators.	Comparison of a range of models from the literature regarding resolution, data sources, etc.	Comparison of national vs. subnational budgets, with and without extensive grasslands.	

on manure management in the EU from National Inventory Reports to the UN Framework Convention on Climate Change (UNFCCC). We also constructed a model to estimate how much biogas could be produced, considering technical and economic constraints such as the minimum production scale and the substrates' dry matter concentration and C:N ratio.

Main findings

With regard to the aim of the paper, we showed that the current EU biogas production could roughly be doubled based on these substrates, and furthermore how the potential depends nonlinearly on the considered constraints.

Significance for this thesis

To estimate the quantities and geographical distribution of substrates, we combined a range of datasets in a novel way. The paper estimates N flows in the substrates, and with minor additions would also produce P flow estimates. The model has roughly 1 km spatial resolution, the most detailed in this thesis.

1.3.2 Paper B

Aim and approach

The aim was to contribute to the development of N balance/footprint indicators for products. We did this by estimating three different N indicators for Swedish conventional and organic milk production. The main data source was a large set of more than 1,800 farm-gate N budgets collected within a Swedish farm advisory project. We used additional data sources to estimate the N flows associated with production of feed purchased to the farms. We carried out some statistical tests and uncertainty analysis to test the robustness of the findings.

Main findings

We concluded that the new indicators calculated for production systems can be useful if the aim is to give environmental information about a product. However, these new indicators are a complement to, not a replacement for, the traditional farm-gate balance, which says something distinct about environmental pollution per unit area. We showed that a potential bias in the estimation of symbiotic nitrogen fixation could possibly switch the ranking of conventional and organic milk production in terms of the three indicators.

Significance for this thesis

The paper is one of few publications that have directly compared N indicators between farm level and production system level. It contributes new knowledge about the technical challenges in estimating nutrient budgets for production systems as compared to individual farms, as well as a detailed discussion of how system boundaries determine the appropriate uses of the indicators. It also highlights the importance of uncertainties about symbiotic N fixation.

1.3.3 Paper C

Aim and approach

The aim was to discuss different perspectives on what the N footprint, a relatively new indicator, can and should be used for. We critically assessed proposed use cases and N footprint models, drawing heavily on the last decade of debate over the carbon and water footprints in and around the Life Cycle Assessment community.

Main findings

We argued that the usefulness of the N footprint must be evaluated in relation to its intended audience and purpose. We argued that the N footprint is appropriate for some, but not all, of its proposed uses. In particular, since the N footprint is a rough proxy for the various environmental impacts that N emissions may cause, it is problematic to use the N footprint in ways that suggest a precise correspondence to environmental impacts. Additionally, our review of N footprint models identified some areas where N footprint models differ and future research can help to improve the relevance and comparability of N footprints.

Significance for this thesis

Among the five papers in the thesis, this is the one that says most about the inherent limitation that agricultural N budgets do not account for N flows and impacts beyond the agricultural system (research question 2a, page 4). Moreover, the detailed comparison of existing N footprint models in the paper provides concrete examples of how limited data and several conceptual issues determine the appropriate uses of the N footprint.

1.3.4 Paper D

Aim and approach

The aim was to investigate how subnational nutrient budgets could complement or replace national nutrient budgets as a tool to monitor environmental risks in the EU. We estimated P budgets for 243 subnational regions in the EU, aiming to use regularly updated international datasets to the extent possible. We also tested a simple method to determine the effect of redrawing the system boundary, as suggested by Eurostat, to exclude extensively managed grasslands.

Main findings

Several data gaps limit the accuracy of these subnational P budgets. We devise methods to fill some of them, in particular for crop production and mineral fertilizer use. Although the estimates have important uncertainties and biases, the results suggest such a large subnational variation in P balances to warrant further data collection and method development in this direction. Furthermore, excluding extensively managed grasslands also makes such a large difference in some regions of the EU that further data collection and method development would be valuable.

Significance for this thesis

The paper draws together the interrelated research questions of how geographical resolution, time coverage of data sources, and system boundaries interact to determine the accuracy and relevance of nutrient budgets and derived indicators.

1.3.5 Paper E

Aim and approach

The aim was to improve data coverage and methods to estimate annual 1961–2013 time series of national cropland N budgets in 26 present-day European countries, building on prior work by Lassaletta et al. [35]. The paper represents a major data fusion effort, drawing on the international FAOSTAT and Eurostat databases as well as many national databases and other publications.

Main findings

The Eurostat crop production database, combined with numerous national data sources, provides a much improved picture of fodder crop production on arable land in Europe. A range of literature sources provide longer and better time series on the division of synthetic N fertilizer application between cropland and permanent grassland. Several other data quality issues are addressed.

Significance for this thesis

The paper explores the data limitations in national agricultural nutrient budgeting for Europe. The paper shows how, with considerable effort, certain limitations of international databases can be overcome using national data sources. The work also showed us that the separation of cropland and permanent grassland is demanding in terms of data, but makes an important difference for the N budgets.

CHAPTER 2

Background

This chapter is divided into four sections. Section 2.1 covers the current state of knowledge about the N and P cycles and the environmental problems caused by N and P pollution. Section 2.2 reviews some facts about agricultural N and P budgets and derived indicators. It describes the construction and main uses of the budgets and how they are estimated. Section 2.3 discusses the use of the nutrient balance as an environmental indicator. Section 2.4 points to some of recent research in nutrient budgeting that this thesis builds on.

Before we continue, I want to comment on some word choices. First, any talk of N flows, unless otherwise noted, refers to reactive N, i.e., any form except dinitrogen (N_2). Second, a recurring theme in this thesis is to distinguish emissions from effects, i.e., to distinguish the release of N and P into the environment from the resulting environmental effects. In this thesis I write about “emissions” to refer to N or P that is released into the environment. I write “losses” almost as a synonym, with the difference that N “losses” from agriculture may include denitrification to N_2 in addition to emissions of reactive N. I write “effects” or “impacts” to refer to the various things that happen in the environment. I use the more neutral-sounding “effect” in general cases; the word “impact” is reserved for unpleasant effects. “Pollution” is a useful but vague word that I use more loosely to refer to an undesirable emission process or to an undesirable state of the environment. Similar semantics are used throughout the appended papers.

2.1 N and P cycles and environmental problems

The remainder of this thesis assumes a basic acquaintance with the different ways that N and P are transformed, transported, and stored in natural and agricultural systems, as well as the main environmental effects that different forms of N and P have in the environment.

However, since my research almost exclusively deals with the construction of agricultural budgets and the use of very simple indicators derived directly from agricultural budgets, I would say—at the risk of sounding careless—that you can understand, and probably could have produced, most of this thesis knowing only the basic facts covered in Chapter 1.

Of course, this is an exaggeration. For a beginning researcher in the field, or anyone else who wants to seriously engage with these issues, it would be ridiculous to not read up on more than the bare minimum. But there is some truth in the previous paragraph. You do not need to know the oxidation numbers of N in different chemical compounds; you do not need to know which types of soil microbes are responsible for different reactions; you do not need to know the lifetime of N_2O in the stratosphere. These things are fascinating to learn about, and doing so will generate many relevant questions, but you can follow this thesis with a fairly rudimentary understanding of the N and P cycles.

While the main motivation of this thesis is the environmental N and P pollution from agriculture, there are also other aspects worth to remember. P fertilizer is produced from mined phosphate rock and thus contributes to a resource scarcity issue, at least in the long run [9, 43]. A related and more immediate problem may be the geopolitical issue that the world's food supply depends on P reserves which are concentrated in a handful of countries [44]. For N fertilizer production, there is no scarcity of materials since four-fifths of the atmosphere is N_2, but it is worth to remember that the Haber-Bosch process uses about 2% of the global energy supply [9] in a time when rapid decarbonization of the energy system is needed.

2.1.1 Relevant literature on different levels

I end this section with a list of relevant literature on various topics and with different degrees of technical detail.

An accessible, mostly jargon-free, and thoroughly referenced overview of the N and P cycles, related environmental concerns, and potential solutions in agriculture of roughly 100 pages is given in Ref. [9]. This would be a good start for many audiences.

The natural processes and basics on environmental issues are covered in many textbooks and review articles. Good textbook treatments of the basics are found in Refs. [45, 46] and [39, 47]. (The first two are somewhat more technical.) For a deeper understanding of agricultural N budgets, it is useful to read more on the terrestrial N cycle, e.g., in Refs. [48, 49]. Some recent findings about (mainly microbial) transformations in the N cycle are reviewed in Ref. [50]. My own summary of the N cycle, global stocks and flows of N, and the many N-related environmental problems, aiming to show the complexity without diving into every rabbit hole, is found in the background chapter of my licentiate thesis [51].

Concerning anthropogenic sources and environmental issues, recent overviews are found in Refs. [6, 10, 39, 52]. A quantitative chronology of the changing N cycle ca. 1860–2000 is given in Ref. [38].

This list would not be complete without mentioning the massive European Nitrogen Assessment [53] from 2011. Its 600 pages and 26 chapters covers terrestrial, aquatic, and atmospheric flows and transformations, environmental effects, and agricultural and policy perspectives. Despite the title, it is relevant reading far beyond Europe.

2.2 Nutrient budgets and balances

Agricultural nutrient budgets have been used for at least 100 years [54]. A main motivation for the early work, apart from general scientific curiosity, was certainly the aim to increase agricultural productivity and make efficient use of limited resources. As environmental concerns grew stronger from around the 1970s, nutrient budgets were increasingly used to assess the environmental risks associated with intensive agriculture [55–58]. Today, as outlined in Chapter 1, nutrient budgets are widely used as a research tool, as a decision support tool in agriculture, and to evaluate policies and monitor environmental risks on national and international level [17–21].

The remainder of this section is divided into three parts. Section 2.2.1 reviews the most common budgeting approaches: How are they defined, and why? Section 2.2.2 looks closer at actual budget calculations: How are they estimated, and what are the main uncertainties that follow? Section 2.2.3 reviews some evidence concerning changes in the nutrient stocks of agricultural soils.

2.2.1 Budget definitions and concepts

A nutrient budget is a list of nutrient inputs and outputs across a system boundary in a given time period. In agricultural budgets, the time period is almost always one year, commonly a calendar year but sometimes the crop year (starting with sowing of winter crops) [40, 59, 60].

The inputs and outputs are measured in elemental nutrient quantities, i.e., kg N or kg P, and usually divided by a reference area so that the units become, e.g., $\mathrm{kg\,N\,ha^{-1}}$ (or $\mathrm{kg\,N\,ha^{-1}\,y^{-1}}$ to be explicit about the time dimension). This is so common that budget terms are sometimes casually discussed in "kilograms", leaving the nutrient, area, and time implicit.

The words input and output require a clarification in relation to N budgets. N budgets are budgets of reactive N. When there is industrial or biological fixation of N_2 or denitrification to N_2 within the system boundary, this is technically not an inflow or outflow in the same sense as something that crosses the system boundary. Nevertheless, it is convenient (and customary) to abstract away from this and list fixation and denitrification as inputs and outputs, respectively.

The balance is defined as balance = inputs − outputs. What does it represent? In one sense, it is simply what is "left over" in the budget. However, I think it is important not to think of the balance merely as a random pile of leftovers. In fact, careful choices have often been made to construct the list of inputs and outputs so that the balance will be useful for certain purposes.

Main types of budgets

Before looking closer at the balance, it will help to introduce the main types of agricultural nutrient budgets. There are many and sometimes confusingly similar types of nutrient budgets. It is not my aim to cover all of them here. The three main types of nutrient budgets that have appeared in my research, and also, according to the Eurostat/OECD handbook on nutrient budgets [60], the main types in use are the following:

- A farm-gate budget, or farm budget, accounts for nutrients entering and leaving a farm (not only through the farm gate, despite the name). The main inputs are typically purchased fertilizers and livestock feeds, manure traded from other farms, and, in the case of N budgets, symbiotic N fixation and atmospheric N deposition. The outputs are the sold crop products, sold animals or animal products, and any manure traded from the farm.
- A soil surface budget, or soil budget, accounts for nutrients added to and removed from the soil. The main inputs included are the applied quantities of synthetic fertilizer and manure, and, in the case of N budgets, symbiotic N fixation and atmospheric deposition. The output is the crop harvested from the field.
- A land budget, or gross nutrient budget, is similar to a soil surface budget, with the main difference that instead of applied manure it includes agriculture's total manure supply, i.e., excreted manure in livestock houses adjusted for any traded manure. The important difference between total manure supply and applied manure is the N that escapes to the environment from animal houses and manure storages, i.e., gaseous losses as ammonia (NH_3), nitrous oxide (N_2O), N_2, and so on. For P budgets, excreted P is practically equal to applied P, and therefore gross P budgets are practically equivalent to soil surface P budgets [60]. The gross nutrient budget is the approach used in the

national Eurostat/OECD budgets, to calculate the indicator known as gross nutrient balance [21].

In addition to the main flows mentioned above, there are some minor flows that are sometimes included in these budget types, such as nutrients in seed and planting materials, field burning of crop residues, atmospheric P deposition, and nonsymbiotic N fixation [60].

Net changes in soil nutrient stocks are sometimes accounted for in budgets. Concretely, this is be done by including a net accumulation term, with the sign convention that soil nutrient stock increases are positive numbers and decreases are negative. The net accumulation can then either be listed among the budget outputs [61], or the budget equation can be adjusted to include the term separately: inputs = outputs + net accumulation + balance [29].

Finally, there are two other budgeting approaches which only tangentially relate to the research in this thesis but are nevertheless relevant to mention:

- The Net Anthropogenic Nitrogen/Phosphorus Input (NANI/NAPI) budget approach aims to include, as the name suggests, the anthropogenic net inputs in a system [42, 62]. On the regional/watershed scale, where the NANI/NAPI approach is most commonly used, the main flows accounted for are synthetic fertilizers and net import of food and feed nutrients, and in the case of NANI, also symbiotic N fixation and atmospheric deposition of oxidized N.[1] The reason to include only oxidized N deposition is that reduced N in the atmosphere mainly originates from ammonia emissions from manure management, which is a recirculation of prior inputs. The oxidized N deposition, in contrast, approximates the NO_x fixed from N_2 primarily in industrial and traffic combustion, and thus represents "new" anthropogenic N input [42].

- Sometimes, farm, soil, and land budgets—and surely other types, too—are estimated with the aim to include *all* the flows, including a complete breakdown of all the different emissions from the system. One motivation for doing so is the long-standing scientific aim to close the budget, i.e., to quantitatively understand where the nutrients are going [54]. Another motivation is that in environmental research, it is much better to know the different forms and pathways for nutrient

[1] Why is the anthropogenic denitrification to N_2 in advanced wastewater treatment not included as a negative contribution? When the NANI approach was devised in the 1990s, this flow was not so big, which might explain why it was not included (Gilles Billen, pers. comm., August 2020). But if the wastewater denitrification would be included, why not also include denitrification in managed wetlands, and maybe even in agricultural soils (ibid.)? See also the appended Paper C for a related discussion on the system boundary between the environment and the non-environment.

emissions than to just know the total quantity [16, 29]. The aim in these estimates—and sometimes a mathematical constraint—is that the balance term should be zero.

This last type of budget which accounts for all the flows is outside the scope of this section and also of this thesis. The remainder of the thesis will focus on the budgeting approaches where only a subset of the flows are accounted for, and the balance term is intended as an indicator.[2] This opens up the next question: What inputs and outputs should be included?

The choice of inputs and outputs to include

What motivates the choice of inputs and outputs to include in a budget? One possible answer is that the inputs and outputs are determined by the system boundaries [16, 60]. While this is certainly true, it does not completely pin down what should be included in the list of inputs and outputs. Here, I will highlight two factors. First, the list of inputs and outputs is established with a more or less explicit aim for what the balance term should represent. Second, inputs and outputs may be excluded on the basis of being unnecessarily complicated to estimate.

The farm, soil, and land budgets to my knowledge have no universal definition that clearly define which inputs and outputs should be included. However, it is clear that the common intention is that the balance at least roughly approximate the total nutrient losses from the system, regardless of whether the sources are anthropogenic or not [16, 60, 61]. The following two examples substantiate this interpretation.

The first example is net changes in soil nutrient stocks, which are sometimes accounted for in nutrient budgets. When they are not, a common interpretation is the one expressed by Oenema et al. [16], that a soil surface balance, for example, "is a measure of the total nutrient loss from the soil, adjusted for possible changes in the storage of nutrients in the soil". A similar interpretation holds for the farm-gate budget, with the difference that the stock change can also be in other compartments on the farm (e.g., stored fertilizers or crop products) [16]. These interpretations are used also in the Eurostat/OECD handbook on nutrient budgets, which is the outcome of discussions among a range of experts [60]. Moreover, the Eurostat/OECD handbook clarifies why soil stock changes should ideally be accounted for, namely that they are reversible. Soil nutrients that one year accumulate

[2] The word "surplus" is mainly used for the balance when some outputs are intentionally excluded so that the balance becomes an indicator of potential environmental pollution [60]. In this thesis, the nutrient balances are usually of this kind (specifically, farm, soil, and land balances), but for consistency I prefer the word balance throughout this thesis. The appended papers mainly use the word surplus.

in soil are not necessarily lost, but can be taken up by a crop in another year. Thus, the accumulated nutrients "can be regarded as a 'useful product' of [the farm] activities". Conceptually, the aim is clearly that the balance approximate the losses.

The handbook's discussion specifically concerns soil budgets, but it is indicated that the same considerations should hold for farm and land budgets. In practice, however, the Eurostat/OECD nutrient budgets do not currently account for soil stock changes since data are lacking [60]. This illustrates how the inclusion of soil stock change is a decision made with reference both to what the balance should represent and to the practical obstacles in form of limited data.

The second example is nonsymbiotic N fixation, which only sometimes is accounted for in farm, soil, and land budgets. While nonsymbiotic fixation can contribute substantial inputs for example in rice and sugarcane (perhaps $30\,\mathrm{kg\,N\,ha^{-1}\,y^{-1}}$), it is probably much smaller ($< 5\,\mathrm{kg\,N\,ha^{-1}\,y^{-1}}$ in many other crops) [63]. However, it is generally agreed that measurement of nonsymbiotic N fixation is exceedingly difficult [63, 64]. In view of these complications, the Eurostat/OECD budget handbook only requires inclusion of symbiotic N fixation, although ideally nonsymbiotic fixation should be included too [60].

In summary, the common intention with farm, soil, and land budgets appears to be to include *all* nutrient inputs crossing the system boundary, and all *intended* nutrient outputs, so that the balance becomes an approximation of *unintended* outputs across the system boundary, which is another way to say nutrient losses. The NANI/NAPI approach instead aims to estimate the net anthropogenic inputs, which could also be expressed as the nutrient excess caused by human activity. However, practical implementations have to differ from these ideal definitions because some flows are impossible or prohibitively complicated to estimate.

2.2.2 Data and estimation methods

The data and methods used to estimate N and P budgets vary depending on system boundaries as well as data availability. It is important to understand how the nutrient inputs and outputs are estimated. The quality of the estimates are heavily influenced by the quality and specificity of available data, and the accuracy of any assumptions or models used in the estimates.

For example, nutrient quantities harvested in crops are normally estimated by multiplying crop harvest data by a nominal N or P content. Crop harvest data may be based on a combination of sample surveys, administrative records, expert estimates, etc., depending on the type and scope of the budget [59, 65]. N or P contents will in best case be measured specifically, but typically are rather assumed or modeled based on literature data [29, 59].

Another example, where the model assumptions are much stronger, is the estimation of symbiotic N fixation in legume crops. Specific field measurements are possible in principle, but since they tend to involve the use of isotope-labeled N compounds, mass spectrometers, and the painstaking work of PhD students [64], the more common approach is to use models generalized from such experimental data [66, 67]. While these models are improving as experimental data accumulate, there is still a considerable unexplained variance which likely depends both on model inadequacies and difficulties in measurement [67].

Uncertainties in inputs and outputs

Like any empirical work, nutrient budgeting involves biases and uncertainties. This is well known in the sense that researchers recognize the risk that estimates are wrong. My impression is that many have a gut feeling for the rough size of their main uncertainties. However, quantitative knowledge of uncertainties and potential biases is lacking. In 2003, Oenema et al. [16] listed and classified potential sources of error in nutrient budgets and noted that "there is little quantitative information about uncertainties in nutrient budgets, though procedures have been developed to analyze uncertainties".

To my knowledge, there is still only a handful of studies that have systematically studied uncertainties in nutrient budget inputs and outputs [68–70]. The overall impression from these studies is that the largest uncertainties are associated with manure, biological N fixation, and roughage fodder. For these flows, errors could easily reach ±20 % [68, 69]. Somewhat smaller but potentially important uncertainties have been emphasized for estimates of crop products, including concentrate feeds [68]. Even smaller uncertainties are those for mineral fertilizer inputs and livestock products [68–70].

2.2.3 Evidence for net changes in soil nutrient stocks

The possibility that soil nutrient stocks change over time is an important issue for the interpretation and environmental relevance of nutrient budgets.

For P, it is clear that large annual soil stock changes can persist over decades. P budgets (NAPI and similar approaches) have shown in several watersheds that the majority of net P inputs may accumulate in soils and the hydrosphere for several decades [71–74]. It is not straightforward to partition this accumulation between soils and the hydrosphere, but it is clear that a substantial share may be stored in agricultural soils [75]. In the EU, it has been suggested that almost five times more P is currently accumulating in soils than is emitted to the hydrosphere [72]; and research in Sweden shows that agricultural soils have accumulated about 700 kg P ha^{-1} over 50 years (about 14 kg P ha^{-1} y^{-1}), corresponding to about 50–100 % of the soil P balance in the same period [76].

This does not mean that large P balances are unproblematic. The river basin budgets merely show that the excess P does not immediately show up at the river mouth. Even the P accumulated in agricultural topsoils, which in principle can be taken up again by crops in later decades [73, 74], does cause downstream pollution, depending on soil type and other factors [76, 77].

Regarding N, the evidence is mixed[3]. On one hand, several budget models for the EU have assumed zero or estimated negative stock changes in EU agricultural soils [29, 61]. Moreover, cropland N balances in several (mainly African) countries have sometimes remained negative for years, which clearly suggests soil N depletion [35]. On the other hand, there are several studies from Europe and North America that suggest net accumulation in agricultural soils [78–81]. In 1999, Smil [82] reviewed several sources of evidence and suggested that while agricultural soils of the world accumulate perhaps 3–6 kg N ha^{-1} y^{-1} on average, soils that receive plenty of N fertilizer may accumulate 25–35 kg N ha^{-1} y^{-1}. In line with this, recent research based on long-term measurements of agricultural soil N in the Mississippi river basin (1957–2010), suggests that soils are currently (1980–2010) accumulating organic N at a rate of perhaps 30–50 kg N ha^{-1} y^{-1} [81]. This would explain about half of the difference between N inputs and outputs in the Mississippi river basin, and it is twice the amount of N that reached the mouth of the Mississippi river in the same period [81]. Although these few indications of substantial accumulation should not be extrapolated to all agricultural soils in regions of high N input, they do suggest that soil N accumulation can play an important part in closing the N budget in agricultural landscapes.

2.3 The nutrient balance as environmental indicator

What speaks for and against the use of the nutrient balance as an environmental indicator? Why is it so widely used? In this section, without aiming for a comprehensive review, I give an overview of some main ideas that shape the current thinking on these questions.

2.3.1 What makes a good indicator?

Indicators are quantitative measures that provide useful information about the status of something important which cannot be directly measured. In

[3] The part of this section concerning N accumulation in soils draws heavily on a section of my licentiate thesis [51].

the wider context of sustainability indicators, Mitchell et al. [83] emphasized that indicators "enable us to gain an understanding of the complex systems around us" and that they do this by: "(1) Synthesizing masses of data; (2) Showing the current position, in relation to desirable states; (3) Demonstrating progress towards goals and objectives; and (4) Communicating current status to users (scientists, policy makers or the public) so that effective management decisions can be taken that lead us towards objectives."

In other words, indicators should not only help us understand something, but also to quantify it in such a way that the present state of a complex system can be meaningfully compared to a desired state. They should also be scientifically sound and practically possible to estimate [41, 83, 84]. These are high ambitions, which of course are only achieved imperfectly. In addition to the inherent difficulty in summarizing a complex system in a simple number, there is also the problem of limited data.

Specifically, van der Werf and Petit [84] categorized a list of 12 proposed indicator methods for agricultural environmental impacts according to whether they focus on inputs (e.g., fertilizer or pesticide use), emissions (e.g., nutrient or pesticide emissions), or resulting system states (e.g., biodiversity or water quality). Broadly, van der Werf and Petit divided indicators into *means-based* and *effect-based* and argued that a major challenge in indicator construction is to navigate the trade-off that means-based indicators tend to be much cheaper to quantify but do not allow an actual evaluation of environmental impacts. A concrete suggestion was that the usefulness of means-based indicators could be empirically evaluated by comparing them to effect-based indicators, to check how well they correspond to one another.

In a similar vein, the official agri-environmental indicators of the EU are categorized according the DPSIR framework (Drivers, Pressures and benefits, State, Impact, and Responses) [85]. Without going into detail, the development and expert evaluation of indicators in these different categories shows again and again the difficulty in identifying indicator definitions that capture something essential without being practically impossible to quantify [26].

2.3.2 The relation between balances and impacts: uncertain and variable

What kind of relation is there between nutrient balances and environmental impacts? Using the simple distinction between means-based and effect-based indicators, nutrient balances are best described as means-based. In the EU's DPSIR nomenclature, the gross nutrient balance is classified among the Pressures [26]. In other words, the connection between balances and environmental impacts is a bit fuzzy. Conceptually, it can be divided into

(1) the connection between balances and losses, and (2) the connection between losses and impacts.

Regarding the connection between balances and losses, Section 2.2.1 explains why the farm, soil, and land budgets can all be interpreted as approximations of nutrient losses from these systems, possibly adjusted for the net nutrient accumulation within the system boundary. The quality of the approximation hinges on correct estimates of inputs and outputs and of the net accumulation. The accuracy of these estimates in general is not known. As mentioned previously, several inputs and outputs could have an error margin of ±20 % or more [68, 69], but clearly this is heavily dependent on the details of the data sources. Regarding net accumulation in soils, Section 2.2.3 shows that on regional or watershed scales it can account for a considerable part of N and P balances, but in specific cases it is difficult to know. The conclusion of all this is that N and P balances, especially those based on the assumption of zero soil accumulation, have to be approached as quite rough approximations of nutrient losses.

The connection between losses and environmental impacts adds yet another layer of uncertainty and variability. The complexity of the N and P cycles imply that the environmental effects are variable, uncertain, and may occur far from the time and place of the original emission [38, 39]. P from agricultural systems can be caught for decades buried in sediments or cycling in freshwater ecosystems before continuing downstream to further freshwater and eventually marine ecosystems (Section 2.2.3). N pollution is arguably even more complex: N can escape from agricultural systems as NH_3 into the atmosphere, as nitrate ($NO_3{}^-$) or dissolved organic N into water ecosystems, as N_2O into the atmosphere; then, through the “N cascade” [7], all these forms can react and recycle several times before eventually being denitrified to harmless N_2. Or, depending on soils and climate, much of the N lost from agriculture can denitrify to dinitrogen already in the agricultural soil [86], in which case a large N balance arguably indicates a resource waste but not an environmental problem. In summary, the environmental impacts that follow from a given quantity of N or P losses is not only uncertain but also highly variable.

At this point, it could sound as if nutrient balances have hardly any relation at all to environmental impacts, but that conclusion would be too pessimistic. After all, research has demonstrated that watershed nutrient balances predict N and P pollution in downstream surface waters [42, 74, 87] and that regional N balances predict $NO_3{}^-$ concentrations in groundwater [88], albeit sometimes with substantial delays. NH_3 emissions enter the environment immediately and thus clearly contribute to N balances, although the NH_3 component cannot be recovered from the balance without additional information. Therefore, the conclusion is merely that the connection between nutrient balances and environmental impacts is uncertain and variable in quality and quantity, and may be distant in time or space.

To summarize, nutrient balances are simple and cheap indicators which have a rather fuzzy connection to environmental impacts. There are several lines of evidence that nutrient balances generally predict environmental impacts, but due to a number of well-known caveats to this general prediction, the nutrient balance is widely recognized as an indicator of *potential* environmental pressures [16, 26, 40, 60, 61, 89].

2.4 Some recent advances in nutrient budgeting

Chapter 1 already mentioned a few specific research strands around agricultural nutrient budgets that underlie the work presented in this thesis. Here I briefly return to these developments as a context for my own work.

The question of geographical resolution in regional nutrient budgets is not new. The possibility of replacing or complementing national nutrient balances with subnational balances has been discussed at least since the 1990s [26, 90]. In the last 15 years, several research projects have estimated subnational N and P balances in Europe on different scales. For example, Grizzetti et al. [27] and Leip et al. [28] have estimated budgets with 1–10 km resolution the EU15 and EU27, respectively; Hong et al. [24, 30] have applied the NANI/NAPI approach to 53 subnational regions in the Baltic Sea catchment; and Le Noë et al. [31] have estimated N and P budgets for 33 subnational regions of France.

These efforts arguably have had two main functions. One is to demonstrate that subnational N and P balances as indicators may reveal many potential nutrient pollution (and depletion) problems that are not seen in national balances. The other is that they help to create a deeper understanding of how agricultural structure—for example the varying degrees of specialization and production intensity—shapes nutrient flows on subnational level [31].

Similar to how increased geographical resolution can reveal previously unseen variation in nutrient balances, there is a question of conceptual system boundaries which can have the same effect: What should be the reference area of a regional nutrient budget? Specifically, it can make a large difference whether nutrient budgets include the whole agricultural area of a region, or exclude permanent grassland, or at least exclude the most extensively managed grassland. For the Eurostat/OECD nutrient budgets, it has been discussed whether budgets should cover only the “potentially fertilised” area to avoid misleading results in countries with large areas of unused or extensively managed land [21, 60]. At present, however, the Eurostat/OECD budgets cover the whole Utilised Agricultural Area (UAA) since clear definitions and data are lacking for the concepts of “potentially fertilised” and “extensive” [21, 60].

Historical time series of N and P budgets also help create a deeper understanding for the relation between agricultural structure, productivity, and nutrient pollution. For example, Bouwman et al. [3] have estimated N and P budgets for global agriculture in 1900, 1950, and 2000; Lassaletta et al. [4, 35] have established national N budgets for cropland (i.e., excluding permanent grassland) covering almost the whole world 1961–2014; and Le Noë et al. [25] have expanded their French subnational N and P budgets to the period 1852–2014. Together, these studies make visible not only a tremendous increase in productivity and pollution pressures, but also that there are huge variations in agricultural structure and productivity between countries and over time.

Various types of nutrient budgets for production systems have been proposed in the last decades. In the early 2000s, Schröder et al. [22] and Bleken et al. [32] noted that farm-gate nutrient balances might be misleading as agricultural and environmental performance indicators if feed production or other important supporting processes occur on other farms, and therefore suggested ways to incorporate information about the combined production system. Several detailed expansions of this idea have followed. For example, Godinot et al. [23] have proposed an indicator called SyNE (system N efficiency); Mu et al. [34] have proposed a “chain” N balance (demonstrated in dairy production systems); and Leach et al. [33] and several others [91, 92] have developed the “N footprint” as an indicator of N emissions caused by products and services.

Two common features of these recent advances is that (1) they expand and explore the possible uses of nutrient budgets and indicators, and (2) they push the limits of what can be done with available data sources. This combination of issues is the topic of this thesis.

CHAPTER 3

Research approach

This chapter is about some methodological issues not covered elsewhere in the thesis. Chapter 2 has already described and motivated the core ideas and methods that are used in the research area and which I have built upon. The appended research papers provide detailed method descriptions. This chapter is divided into four sections. Section 3.1 shares a few thoughts about mathematical models as vehicles for scientific inquiry. Section 3.2 lists the main data sources used in the appended papers. Section 3.3 then explains and motivates my approach to data cleaning and data analysis, a central but partly hidden part of my research. Section 3.4 ends with some mundane but important concerns around the computer software that we build and use in our work.

3.1 On models and modeling

One of my co-authors recently asked, with reference to a manuscript draft of Paper E: "What is a model?" This question started an interesting conversation. Partly it turned out to be a question of nomenclature, but partly it was about something more profound. I think that it is reasonable to ask whether Paper E presents a model or just a collection of data. Without getting too philosophical, I will therefore try to say something about what I think a model is and how that matters in our day-to-day work.

3.1.1 Studying reality, or the model?

Here are four descriptions of the work behind this thesis: I have been …

- studying nutrient flows in agriculture.
- studying models of nutrient flows in agriculture.
- creating models of nutrient flows in agriculture.
- studying how to create models of nutrient flows in agriculture.

The peculiar thing is that all this has been ongoing at the same time. To work with a quantitative model based on empirical data is to constantly climb up and down a ladder of abstraction, revising beliefs about reality, beliefs about the model, and beliefs about the data and processes that generated the data. The model construction tends to get intertwined with model validation and analysis of results.

From an epistemological standpoint this creates several difficult questions. Are we learning about the model or about reality? Are we adjusting beliefs about reality using the model and the data, or are we adjusting the model and the data to our beliefs about reality? Is the research method scientific in the sense that it will "respond" to reality so that hypotheses can really be falsified?

3.1.2 Representation and idealization

A crucial function of models is that in some way they *represent* a target system. We hope to learn about the target system by studying the model. Practitioners and theorists agree that the degree and type of representation can vary enormously, from statistical models which mainly describe a pattern seen in data, to mechanistic models which are thought to have some correspondence to the "actual process" [93]. Expectations on a model to explain and create understanding have to be adjusted accordingly [93, 94].

Models are usually *idealized* representations of reality. Idealizations can be categorized as Aristotelian or Galilean [94]. Aristotelian idealization is to disregard certain aspects of the target system to focus on what is relevant. For example, a nutrient budget describes one aspect of a system, disregarding most of reality without necessarily distorting it. Galilean idealization is to consciously distort reality. For example, we have assumed a uniform protein content in wheat across all time and space although we know this is not true. Real-world models usually contain both types of idealizations, but it can be helpful to distinguish them conceptually.

3.1.3 The nutrient budget is a model

In my view, an agricultural nutrient budget is a model. In its most basic form, it consists of a system boundary and a list of nutrient inputs and outputs across that boundary. This is an Aristotelian idealization of the very best kind, facilitating quantitative and distraction-free study of a certain aspect of the target system. In a limited but profound way it represents the target system. Actually so well that I sometimes confuse the model with reality. The budget has an intuitive relation to the things and processes it represents.

In practice, nutrient budgets invariably also require Galilean idealizations.

We estimate inputs and outputs from available data using all sorts of assumptions as explained in Chapter 2. Moreover, the input data usually have to be cleaned and filtered and adjusted in various ways before they are ready to be used. In other words, the data we use are typically not the original "raw" data but an idealized version, sometimes called a "model of data" [94].

3.1.4 Keep the idealizations in mind

Why ask these abstract questions about such a simple model? I think that it is important for simple and complicated models alike. Maybe even specifically in this case, because the nutrient budget has such strong and intuitive representational quality that the model can easily be mistaken for the real thing.

When faced with an unexpected model result I can be tempted to jump to equally unexpected conclusions about the part of reality that the model represents. But an alternative view is that unexpected or unreasonable model results indicate possible errors in the model. By carefully thinking of each model result as a transformation of input data under certain assumptions, it can be figured out which are the most suspect data or assumptions. The reaction is then to test an alternative assumption or data source and check how that affects the results, and so on.

This incremental and somewhat disorganized conversation between beliefs, hypotheses, and data is in my view a central part of what is called "modeling". It is often how I simultaneously validate a model, learn about the model as such, and also learn about reality. In the research behind this thesis, I have rarely validated models through explicit, planned tests of model results against independent data sources. In fact, much of the modeling work we have done has aimed to estimate various quantities for which *no* data are available, and therefore it has often been impossible to perform such direct tests.

This is a strong motivation for keeping it simple. Small, simple models are easier to remember, easier to explain, and easier to understand and validate. Identifying the assumptions or data sources responsible for a particular result is *much* easier to do well with a small model.

The idealizations built into a model create limitations to what it can be meaningfully used for. Remember and mentally account for the idealizations and you will be fine. Disregard them and you risk making terrible mistakes.

3.2 Data sources

This thesis builds entirely on secondary data from various international and national databases and literature sources. The international databases used in the different papers are the following.

- Papers A, D, and E use Eurostat's agricultural and environmental statistics, including: annual crop statistics [65, 95], the many data tables of the Farm Structure Survey [96], and fertilizer sales and consumption [97, 98]. In additional to national data, we have used subnational data mainly on the NUTS2 level, i.e., the second subnational level of Eurostat's hierarchical regional classification system NUTS. Papers A and D also use spatial data about the NUTS regions [99].
- Paper A additionally uses spatially explicit data from the Corine Land Cover database [100] and from FAO's Gridded Livestock of the World v2 [101].
- Papers A and E use data on manure management systems from the National Inventory Reports (NIRs) submitted to the UN Framework Convention on Climate Change (UNFCCC) [102].
- Paper D uses data on fertilizer use from the Farm Accountancy Data Network (FADN) [103] as well as spatial data on the FADN regions [104].
- Paper D also uses an international collection of manure excretion coefficients for livestock in the EU countries [105].
- Papers A, B, D, and E use data on nutritional and other properties of livestock feeds from Feedipedia [106].
- Paper E uses data from the FAOSTAT database [107], not only by building on previous results by Lassaletta et al. [35], but also by making new analyses of FAOSTAT's statistics on crop production [108, 109], fertilizer consumption [110, 111], and land use [112].
- Paper E also uses data from the International Fertilizer Association's IFASTAT database on fertilizer consumption [113].

National data sources were mainly used in Papers B and E. Paper B builds almost exclusively on national data, most importantly a database of farm-gate nutrient budgets from the Swedish farm advisory project Focus on Nutrients (Greppa Näringen). Paper B also uses national statistics on crop production and fertilizer use. Paper E uses a range of databases and reports from national statistical agencies (in addition to other literature sources) to fill data gaps in the international databases mentioned above.

3.3 Dealing with data

I have spent a large share of my doctoral studies collecting, understanding, selecting, and adapting data from these data sources for various modeling purposes.

It is perfectly normal for datasets, even highly structured databases, to contain errors, irregularities, missing values, and other issues that need to be resolved before the data can be used. Necessary tasks may include, for example, to:

- parse unstructured data,
- transform to compatible units of measure,
- select relevant subsets,
- aggregate to groups other than the original observation units,
- identify and remove or correct errors,
- identify and fill data gaps, and
- rescale or make other adjustments for agreement with related datasets.

I collectively call these tasks *data cleaning*.

Although every data source requires specific solutions, I have through these years maintained and refined the following loose set of working principles for collecting and cleaning data.

- Read the documentation. Statistical databases and major scientific datasets usually have technical documentation that give important information about how the dataset was created. This gives hints to how it can be used and what might go wrong.
- Check for consistency and completeness. Even high-quality databases are subject to human and machine error which can cause severe inconsistencies or data gaps. Try to think of criteria that should be fulfilled (e.g., the sum of each row should be 100 %, or all 28 EU countries should be covered) and check for them. Do not blindly assume that datasets conform to their specifications.
- Use graphical tools when possible. Humans are extremely strong in visual pattern recognition. By inspecting a clever plot, you learn more about the data quality in seconds than by browsing data tables for hours.
- Data clean you must. It is not more correct or “objective” to leave datasets untouched than to judiciously remove errors or fill data gaps.

The "raw" data are not the truth but simply the state in which the data publisher left them.

- Prefer statistical surveys to expert estimates. Experts are humans too, and thus subject to the same cognitive biases as everyone else. Expert estimates also tend to be less precisely defined and documented than statistical surveys.
- Do not fear expert estimates. When there are no other data, or when the other data have inadequate quality or coverage, expert estimates (including your own) may be useful and even necessary. When no hard facts are available it may seem more objective to assume a uniform distribution than to rely on guesstimates. But the uniform distribution is not in general more objective than anything else, and should therefore be used only if there is a specific reason to believe in it [114, 115].
- Clean with a purpose. This is admittedly more a lesson learned than a principle I brought into the research project. It is usually both unpractical and unnecessary to clean datasets to perfection. The aim of data cleaning is not to rectify all data quality issues, but to identify and deal with them in proportion to their severity.

3.4 On scientific computer software

A small but growing share of scientific publications in data-driven and computational science adheres to the loosely defined standard known as reproducible research. As explained by Peng [116]:

> The standard of reproducibility calls for the data and the computer code used to analyze the data be made available to others. This standard falls short of full replication because the same data are analyzed again, rather than analyzing independently collected data. However, under this standard, limited exploration of the data and the analysis code is possible and may be sufficient to verify the quality of the scientific claims.

This idea resonates with me, at least in theory. Reproducible computations facilitate external verification, validation, and derivative works, which are cornerstones of successful science. In practice, though, implementing and sharing fully reproducible computations is time consuming and surprisingly difficult. How to ensure that the program will run on someone else's computer? How to ensure that it will even run on my own computer in ten years [117]? Thankfully, gradual progress is possible. Peng [116] argued that a decent starting point is to share the source code, even if it is disorganized and poorly documented. A next step is to clean up and document the source code, and yet another to share input data and other information needed to actually run the code.

I have at least taken some steps in this direction. With Paper A, we shared the source code and instructions for how to obtain the data from various places [118]. The only person who heroically tried to run the code (to build upon the results in another model) got stuck because the input data were no longer available at the same address or in the same version. Luckily I still had the data and we could rerun the program. Lesson learned: Statistical databases change over time and it is therefore necessary to save the original data. I have since built a software tool to download, archive, and read timestamped tables from Eurostat's database [119]. In Paper E I have finally taken the step to specify which exact data versions we used.

With Paper B we did not share source code. We were not allowed to share the data, and I therefore incorrectly reasoned that sharing the source code would be pointless. With hindsight, we should have shared the source code because it says a lot even without the data [116].

With Paper D we published cleaned-up and commented source code for the whole calculation. We archived both the source code and the model output data in the research data repository Zenodo [120].

I agree with Peng [116]: "Perhaps the biggest barrier to reproducible research is the lack of a deeply ingrained culture that simply requires reproducibility for all scientific claims." If such a culture becomes reality, people like me will have to spend more time documenting their work, but it would be so worth it.

In this context, I also want to acknowledge some of the open source softwares that have made this thesis possible in its present form: the Python [121] and R languages [122]; the SciPy ecosystem (SciPy, NumPy, pandas, matplotlib, and scikit-learn) [123–127]; Git [128]; and Zotero [129].

CHAPTER 4

Main findings and reflections

Chapter 1 outlines how the appended papers help to answer the research questions of this thesis. This chapter adds further details and reflections to that description. This chapter is organized around the research questions posed in Chapter 1. Section 4.1 is about how available data and methods limit the possibility to estimate European N and P budgets at various degrees of geographical resolution, with different system boundaries, and in different time periods (RQ1). Section 4.2 is about the inherent limitations of agricultural nutrient budgets following from the fact that they do not account for what happens in beyond the farm (RQ2a). Section 4.3 is about the limitations of agricultural nutrient budgets following from uncertainties and lack of data (RQ2b).

4.1 Limits to accuracy and level of detail (RQ1)

Each of the appended papers say something about the accuracy and level of detail that can be attained in agricultural nutrient budgets. Most importantly, the papers A, B, D, and E explore this question by estimating nutrient budgets or parts of nutrient budgets. Most of our work has been on N budgets—Paper D is the only exception—but given the many similarities between estimation methods for N and P budgets, the results presented below largely apply to both types.

4.1.1 International databases go a long way

As explained in Section 3.2, our work has used both international and national datasets. My belief is that we have been fairly successful in exploring the limits of what can be achieved using the international datasets. However, when expanding the view to national datasets, there is much more to discover. For example, the study of N flows in Swedish organic and conventional milk production systems in Paper B demonstrates that national data sources can far exceed the international databases in detail. In view of the details presented in our Paper B, in view of the French subnational 160-year history presented by [25], and in view of the many other data sources I have heard about or seen, it is exceedingly clear that we have not reached the depths of what can be achieved using national data sources. The work we present in Paper E makes use of national data sources, but it is a scouting mission rather than an inventory of available data. In summary, while we have also felt the limitations of national data supply, the conclusions presented here are primarily about international data sources.

It can hardly be overstated how valuable the international datasets listed in Section 3.2 are for research on agricultural production systems. Individually, resources like the FAOSTAT, Eurostat, FADN, IFASTAT, and Corine Land Cover databases are valuable because they provide easy access to harmonized data on key variables in agriculture. In combination, they enable analyses that would otherwise be practically impossible. These international datasets are public goods that deserve praise and continued investment.

4.1.2 Crop areas and harvests

The most complete, consistent, and accurate of these datasets are arguably those about commodity crops and livestock. There is a long tradition of more or less standardized weights and measures for all the major commodity crops, including cereals, oilseeds, roots and tubers, etc., as well as the main types of production livestock. Production of meat, milk, and eggs has also been recorded in mostly comparable ways for decades. The FAOSTAT database covers most countries of the world since 1961.

The Eurostat database typically has data on its member states since around the time of their accession to the EU. For the earliest members of the EEC, agricultural statistics are available since the 1950s. The subnational data have become increasingly complete in the last decades.

Eurostat's Farm Structure Survey is notable for its completeness and internal consistency, also on subnational level. We made heavy use of Farm Structure Survey data in Papers A and D. The Farm Structure Survey does not report crop harvests, only areas, and therefore we used a combination of data from the Farm Structure Survey and the annual crop statistics. A detail worth to highlight about the Farm Structure Survey is that its crop areas follow the "main area" definition, which counts every field only once. In contrast, the annual crop statistics reports sown or harvested areas and may therefore due to double- or intercropping sum to more than the total cropland area. Detailed comparison between the two data sources is therefore complicated. However, since most cropland in the EU is harvested only once per year it is an acceptable approximation for many purposes to ignore the difference between harvested and main areas.

Our only work with higher resolution than Eurostat's NUTS2 level is presented in Paper A. We disaggregated subnational crop statistics using the Corine Land Cover dataset using the idealization that each grid cell defined as arable land contains the same crop mix as the enclosing region. This is idealization is not straightforward to validate. However, data supply is improving on finer scales as a growing number of models of spatial crop distributions have started to use advanced remote sensing techniques in combination with agricultural statistics [130]. These are promising alternatives to our simple model but fall outside the scope of this thesis.

Some important exceptions to Eurostat's mostly complete and consistent crop production data are found for arable fodder crops such as green maize, temporary grassland, forage legumes, and fodder roots. We describe these issues in detail in Paper E. However, note that (see Paper E) the completeness and internal consistency of fodder crop data has improved markedly in the last decade. Today, most EU countries have complete and consistent annual records of fodder crop areas and harvests, also on subnational level (see Paper D).

In Paper E, we present a gap-filled dataset of harvested areas and N quantities in 14 arable and permanent crop categories (excluding permanent grassland) for 26 present-day countries in the period 1961–2013. All but one of the categories is based primarily on FAOSTAT's crop production database. The last category is arable fodder crops, and it is based on Eurostat's annual crop statistics and a range of other data sources.

In Paper D, we present a gap-filled dataset of P harvests in 17 main crop types (including permanent grassland) in 243 subnational regions of the EU28 in year 2013. The gap-filling procedure uses a mix of data from the Farm Structure Survey, the annual crop statistics, and the permanent grassland productivity estimate of Smit et al. [131].

Permanent grasslands, which cover about one-third of Europe's agricultural area, are described with considerably less detail and consistency than cropland. Permanent grasslands are known to vary enormously in management and productivity, but due to lacking data they are often crudely represented in nutrient budgets and other models of agricultural production. The most detailed and comprehensive estimate of permanent grassland productivity in Europe, by Smit et al. [131], is based on a mix of national and subnational statistics, literature data, and expert estimates from the early 2000s. This is the estimate we used in Paper D. There is no regularly updated European database of permanent grassland productivity. Interestingly, the annual gross nutrient budgets established by the EU member states do include permanent grassland, meaning that national experts regularly make national estimates of the N and P harvested and grazed from their permanent grasslands. To my knowledge, there is no standardized reporting of the precise methods and data used for these estimates. If this is not simply a reflection of my ignorance, it would be valuable to study what methods and data are nationally used for these estimates.

The area of permanent grassland reported in European countries sometimes differs between the Eurostat and FAOSTAT databases by considerable amounts and for reasons that are not entirely clear. In Paper D we used the Eurostat permanent grassland areas for consistency with the other Eurostat data sources. In Paper E, the choice was not straightforward since we there used a mix of FAOSTAT and Eurostat data. We finally chose to use the FAOSTAT permanent grassland areas for consistency with earlier work by Lassaletta et al. [35], but it is yet unclear to us what is the best choice in this case.

4.1.3 Synthetic fertilizer use by crop and by region

One of the ways that management of permanent grassland varies is the rate of fertilizer application. In Paper E we build on previous work by Lassaletta et al. [35] and present an improved estimate of how synthetic N fertilizer application has been divided between permanent grassland and cropland in 26 present-day European countries in the years 1961–2013. While I am the first to admit that this dataset is subject to a host of data quality issues, I would say it is the most comprehensive review made of this topic and a substantial improvement compared to the previous state of knowledge. Further improvement of our dataset would probably require the assistance of national experts.

A by-product of our work with the historical division of N fertilizer between cropland and permanent grassland is a compilation of data which during the 1990s and early 2000s also states fertilizer use for the main crop types in many European countries. This data collection (to be published along with Paper E) may prove useful in future research on nutrient budgets and related topics.

A related question is how much N and P fertilizer is used nationally and in subnational regions. On national level, there are annual time series of N and P fertilizer use published by Eurostat, FAOSTAT, and IFASTAT. Eurostat has two different estimates, one based on annual sales according to Fertilizers Europe, and one based on reports from national statistics offices to Eurostat. This makes a total of four different data sources (see Section 3.2). The estimates are variously based on sample surveys, sales data, balances of production and trade data, or a combination of these approaches. As we show in Paper E, these different datasets for N fertilizer broadly agree in most countries and most years, but there are sometimes considerable differences between them.

In Paper E we present a combined dataset building mainly on IFASTAT data, with a few gaps and suspected errors filled from other data sources.

On subnational level in the EU, we have found two international datasets on fertilizer use. One is Eurostat's subnational fertilizer use statistics, which have increasing but still partial coverage of the EU. In Paper D, the Eurostat database provided subnational P fertilizer statistics for 15 of the EU28 countries in year 2013. For the remaining 13 EU28 countries we estimated the subnational distribution using data from the FADN database, rescaled to agree with national totals according to Eurostat. The FADN database reports fertilizer use starting in year 2014, reaching full coverage of the EU28 in 2017. Two complications are worth to highlight here (see Paper D for details). The first is that Eurostat's subnational regions according to the NUTS system does not fully coincide with FADN regions. The second is that the national total quantities do not fully agree between the Eurostat and FADN data. We resolved these issues by projecting the FADN quantities onto the NUTS regions and rescaling them to agree with Eurostat's national totals. We validated the method by comparing the results to Eurostat's subnational statistics in the 15 countries where both data sources were available. The results are not entirely satisfactory. Although the FADN based estimate broadly agrees with subnational statistics, in some subnational regions the quantities disagree by some 5–10 kg P ha^{-1} y^{-1} (averaged over the agricultural area). The reasons for the discrepancies are unknown to us. In summary, subnational fertilizer use statistics are increasingly available in the EU, but there are some remaining data quality issues. Further efforts to validate and combine different data sources would be valuable.

4.1.4 Manure

The nutrient flows in excreted, stored, and applied manure are subject to considerable uncertainties. The N and P quantities excreted by livestock are typically estimated rather than measured. Gaseous losses of N from livestock houses and manure management systems are also typically estimated rather than measured. These gaseous losses have been subject to much research, resulting in estimation methods that are increasingly detailed and standardized [132, 133], but uncertainties nevertheless remain. Additional uncertainties pertain to manure management systems, including grazing management, which vary between livestock classes, over time, and between countries. Standardized data on manure management systems are lacking. The research network RAMIRAN has launched a project to produce Country Manure Management Profiles [134]. However, this work is progressing slowly due to lack of data (Harald Menzi, pers. comm., April 2020).

This thesis makes some minor contributions to improved modeling of manure nutrient flows in Europe. In Papers A and E, we estimate the distribution of manure N flows between different manure management systems using data from the UNFCCC NIRs for the EU28 countries. The exact methods used in each of the NIRs vary since it is up to a team of national experts to establish these estimates according to the reporting guidelines. Thus, our approach leverages the work of national experts without needing to make detailed investigations into each country. For Paper E we used this approach to establish annual time series on manure management systems for the 26 European countries in the paper starting in year 1990. Before 1990, we have not found any international datasets on manure management.

For Paper E we briefly looked into the possibility of using data on manure flows from the annual reporting to the UN Convention on Long-Range Transboundary Pollution (CLRTAP). In particular, the CLRTAP parties' Informative Inventory Reports (IIRs) [135] sometimes contain much information on excreted manure quantities as well as on manure management systems. However, the IIRs are not in a standardized or machine readable format, and the details on manure flows are not part of the standardized activity data published by CLRTAP [136].

On this note, it has been highlighted by experts working with Eurostat's gross nutrient budgets and the annual reporting to CLRTAP and UNFCCC that these reporting duties in some parts require similar data and modeling efforts, and that by harmonizing input data this work could be simplified [60, 137]. If such harmonized datasets could be produced based on national data and expertise, and preferably published in machine readable formats, they would also be highly valuable for other purposes.

4.1.5 Symbiotic and nonsymbiotic N fixation

As noted in Section 2.2.1, biological N fixation should ideally be included in farm, land, and soil N budgets regardless of whether it is symbiotic or nonsymbiotic. However, nonsymbiotic fixation is in many cases ignored because it is generally believed to make only a small contribution in agricultural soils. This thesis acknowledges the uncertainty in nonsymbiotic fixation but makes no progress on this matter.

Concerning symbiotic N fixation, Papers B and E make some contributions. In Paper B, symbiotic fixation was identified as one of the major uncertainties. Moreover, since the organic farms in Paper B had higher legume content in their fodder production, a general bias in fixation estimates would affect conventional and organic milk systems differently. Our sensitivity analysis demonstrated that a possible bias in symbiotic fixation estimates could heavily affect the study outcomes, even shift the order of conventional and organic milk in terms of the three indicators.

In Paper E, the symbiotic fixation is one result calculated using our dataset of arable fodder crop cultivation in the period 1961–2013. We took specific care in the construction of the fodder crop dataset to distinguish between forage legumes, legume/non-legume mixtures, and non-legumes, in order to enable the best possible estimate of symbiotic fixation in cropland. The resulting dataset (to be published with Paper E) shows, for example, that in 1961 the input of symbiotic N fixation in European cropland roughly equaled the input of synthetic N fertilizer, but since then has fallen by 50 %. This development is well-known in principle but quantitatively has hitherto been unclear because no complete dataset of forage legume production has been available.

4.1.6 Estimating budgets for production systems

N budgets for products or production systems are discussed in both Paper B and Paper C. The N footprint, which is the topic of Paper C, is technically and conceptually different from the chain N balance estimated in Paper B. However, the two indicators are very similar in terms of modeling and data requirements, why I make no further distinction between them here.

The model presented in Paper B estimates N budgets attached to the average kilogram of organic and conventional milk. This is achieved by modeling the production system behind the feed purchased to the dairy farms. The feed is co-produced with other crop-based products in the food and feed industry, which in turn purchases crop products from other farms. In our model, the feed user (the dairy farm) is allocated a quantity of that "upstream" crop cultivation in proportion to the economic value of the feed compared to its co-products. The chain N budget is the farm-gate budget of the dairy farms plus the soil surface budget of the upstream feed crop cultivation. The last step is to allocate the resulting chain N budget to the dairy farms' products (milk, meat, and crop products) in proportion to their N content.

This creates several new requirements for data. What feed products are purchased by the dairy farms? What are the co-products of those feed products? How are the upstream crops cultivated? What are the prices used for allocation? What selection of dairy farms should be used to represent the more abstract system of Swedish milk production?

In Paper C, one part of our approach was to compare different models used to estimate N footprints, to better understand the methods and data sources that can be used. It is outside the scope of this thesis to draw any definitive conclusions about the data requirements associated with different methods, but broadly I think that Papers B and C together illustrate two things: (1) it is possible to make reasonable estimates of N budgets for production systems, but (2) the accuracy and consistency of such estimates are severely limited by data supply. Budgets for products or production systems require much more simplifying assumptions than the more traditional farm-gate and regional budget approaches do.

Specifically, it is worth to highlight the importance of international trade in food and feed commodities as an important uncertainty and a rich source of exciting research questions. Our review in Paper C shows that methods for handling international trade ranges from largely ignoring the issue—like we did in Paper B—to using sophisticated economic models of global trade in various commodities. One motivation for the more advanced trade models is that some N footprint models also include NO_x and other N pollutants associated with non-food consumption. But in principle the advanced trade models are appropriate for food consumption too, since food and feed are indirectly connected to markets of fiber, fuel, forestry, fishery, and eventually the whole global economy [138, 139]. This is an area of research where many empirical and conceptual issues are yet to be resolved.

4.1.7 Concluding remarks

N and P budgets have been established for an array of agricultural systems with different geographical resolution, with different system boundaries, and in different time periods. Missing data are routinely substituted by extrapolation from similar systems, deduction from indirect evidence, expert estimates, and so on. Therefore, the question answered here is not so much *whether* as *how* and *how well* budgets can be established for different agricultural systems.

A banal but important observation is that budgets are easier to estimate if the system boundaries correspond to administrative boundaries. For example, since fodder crops and manure mostly circulate within farms, they are not weighed, analyzed, and documented with the same rigor as mineral fertilizers or commodity crops. The resulting uncertainties about fodder crops and manure have a stronger effect on soil and land budgets than on farm-gate budgets. Similarly, national or regional nutrient budgets are much, much easier to estimate accurately if they include the whole agricultural area instead of excluding, for example, permanent grassland (Paper E) or extensively managed grassland (Paper D).

Our work with subnational regions in the EU (Papers A and D) demonstrate that soil and land budgets can be estimated in 200–300 EU regions (somewhat depending on year and other details). These subnational budgets have more data quality issues than national budgets, though. One particular point is that we have hardly found any information about within-country transport of manure. We know that this heavily affects our P budget estimates in the

Netherlands, but probably there are other regions too which we are unaware of. Another point is the partial coverage of subnational fertilizer statistics in the Eurostat database. Data coverage has increased markedly in the last years and it is our hope that these efforts will continue.

Regarding the time aspect, 1961 is an important year because it marks the start of the FAOSTAT database. Before 1961, harmonized international data on agricultural production are very scarce. Eurostat's national agricultural statistics begin in the 1950s, but only for a small subset of the present-day EU countries. Subnational data coverage starts in the 1970s. The countries covered by the Eurostat database has then grown along with the accession and entry of new member states. An important detail (Paper E) is that production of most arable fodder crops is not covered by the FAOSTAT database. Therefore, completing the record for the later EU member states back into the 1960s requires consultation of national data sources.

Finally, I want to emphasize that the "lack of data" discussed in this thesis is sometimes a practical rather than a total lack of data. For example, our work with Paper E has demonstrated that much relevant data is to be found in national databases, papers and reports, and statistical yearbooks, if only enough time can be devoted to finding and collating them. Paper B also shows that, even if there is a lack of data about Swedish milk production in the public library, there is an extraordinary amount of data available in the database of the Focus on Nutrients project. Thus, "lack of data" in many cases means lack of *easily available* data. If the aim is to improve the data supply, this is a crucial distinction to make.

4.2 Inherent limitations of agricultural budgets (RQ2a)

Paper C considers the inherent limitation that the N footprint does not account for actual environmental effects. However, most of the arguments advanced on this point in Paper C would apply equally to an imagined P footprint, and also to other indicators derived directly from agricultural nutrient budgets, such as the chain nutrient balance.

We argue in Paper C that it should not be blindly assumed that the N footprint is an adequate proxy for the impacts that people care about. This has to be evaluated in relation to the intended purpose. We argue that the N footprint is quite good enough for some proposed uses but not for others. We are open for different evaluations, though, because in the end there are no objective criteria. The important point is that there is a need to consider how the N footprint will be understood in the context it is meant to be used.

Can the limitations of the N footprint be overcome by modeling environmental effects? Yes and no. In Paper C we suggest that the N footprint for some audiences could give more nuanced and relevant information if it is disaggregated into different chemical forms or by location of the N losses, as some authors have already done. Emissions of NO_x and N_2O, for example, cause so different impacts that a moderately knowledgeable audience could make use of the extra information. This would be a step in the direction of impact assessment. Many further steps could then be taken by adding increasingly sophisticated models for impact assessment. As we argue in Paper C, though, there is no magical point at which the N footprint would suddenly become a perfect impact indicator, because there is no such thing.

My conclusion is that there is nothing special about the boundary between the environment and the non-environment that determines the adequacy of the N footprint or any other quantitative indicator.

4.3 Limitations of budgets due to uncertainties and lack of data (RQ2b)

The research question is *not* the following: Would nutrient budgets become more useful if we could have any data we wished for? That question has the resounding answer “yes”. The question is: *How* are present and proposed uses limited by uncertainties and lack of data? That question is much more difficult to answer.

To answer this question, this thesis engages with some suggested and demonstrated uses of agricultural N and P budgets that are possible but difficult, and where real improvements can actually be foreseen. In particular, this thesis looks into four contemporary developments in nutrient budgeting (see Section 2.4) that are limited by data.

First, establishing nutrient budgets for products or production systems is the topic of Papers B and C. At the price of some simplifying assumptions (Section 4.1.6) it possible to make such estimates, which are clearly useful for some purposes (see Paper C). One indication of how uncertainties can limit these uses, however, is seen in Paper B: The uncertainty in symbiotic N fixation is large enough to potentially reverse the order of the conventional and organic milk in terms of the three indicator values. This supports the point which we indicated in Paper C, that detailed comparison within product groups with similar N footprint values is difficult. The most obvious value of the N footprint is to illustrate the large order-of-magnitude differences that exist between different types of products.

Second, increasing the geographical resolution of national nutrient budgets is the topic of Paper D. As noted above, the subnational P budgets do have some additional data quality issues. Can they be solved? I would speculate that manure trading will be a substantial data gap for a while yet. In contrast, subnational fertilizer statistics are steadily accumulating, and may in fact largely be available in national databases already. In any case, we argue in Paper D that the amount of additional heterogeneity revealed by the subnational P budgets is so substantial that the higher uncertainty may be acceptable, depending on the purpose. The choice is a trade-off: the national budgets are easier to get right, but they might inadvertently hide much of the subnational variation that may cause problems.[1]

Third, the benefits and challenges associated with different reference areas for national or regional budgets are explored in Papers D and E. In Paper D, one of the research questions was what would be the likely effect of excluding extensively managed grassland. In Paper E, the choice of cropland (excluding permanent grassland) as reference area requires considerable efforts estimating the division of fertilizer and manure inputs between permanent grassland and cropland. Both papers demonstrate that the choice of reference area can have a very important effect on national and subnational nutrient budgets. Both papers also demonstrate that it is much easier to establish accurate national and subnational nutrient budgets if the whole agricultural area is included. The most appropriate system boundary, as always, depends on the purpose of the study.

Fourth, the time period covered by different datasets (see Section 4.1.7) has a strong effect on what can be learned about past and present nutrient flows in agriculture. Specifically, Paper E demonstrates that data on arable fodder crops is incompletely covered by international databases but that national databases contain much additional information. Paper E also shows that the documentation of fertilization of permanent grassland is particularly weak before the 1990s, even if national data sources are consulted. Paper D focuses on present and future data coverage in its explicit aim to use data sources that are now continually updated. A related question which we have not explored in detail, is how well N and P budgets could be estimated in earlier years using Eurostat's subnational data.

In summary, these are four areas of development that are all exciting and valuable, but which are all limited by data scarcity. This thesis shows that the data scarcity seems total on some topics, while on other topics data can be found given resources to consult national data sources. The thesis also shows that creative combinations of existing data sources can be used to fill data gaps. Further exploration of existing national databases, statistical reports, and research literature will surely yield new datasets.

[1] In Paper D, I am responsible for a confusing use of the words accuracy and precision. I meant something like bias and variance, but even that is not a great analogy. If you are a statistician, I hope you can forgive me. However, I think it is still fairly clear what is meant in the paper, so it is best that you just read it and try to ignore those words. The point is simply what I write in this thesis: the national budgets estimates are technically more correct but will tend to hide the variation that we are looking for.

Chapter 5

Closing words

This thesis aims to explore and expand the limits of how agricultural N and P budgets can be used in environmental research and decision-making. The appended papers contribute to this aim by estimating N and P budgets for various systems and by reflecting on what the budgets and derived indicators can be meaningfully used for.

Lack of data has been a central theme in my research. I have actively been looking for the limits of what can be achieved using existing data sources. An understanding that has deepened along the way is that "lack of data" can mean quite different things. Sometimes there is a total absence of useful information. Sometimes there are only expert estimates. Sometimes there are data, but they are incomplete or inconsistent. Sometimes there are data, but they are not available to the public. Sometimes there are data, but they have to be collected from national databases or reports in foreign languages. Sometimes there are excellent data, conforming to well-documented standards, complete and consistent, machine readable, publicly archived.

There is much that can be learned by combining different datasets. The research presented in this thesis builds on a tradition of quantitative modeling that combines statistics databases, output data from scientific models, literature data, and expert estimates. We have devised and evaluated several new ways to combine existing data sources, with varying degree of success. Many exciting and useful insights are yet to be generated in this way.

This thesis investigates how nutrient budgets and derived indicators can and should be used given their various limitations. How to calculate and use footprint-style indicators for agricultural products? How to establish subnational N and P budgets? Can cropland be separated from permanent grassland? I hope that this thesis has deepened the understanding of how the answer to these questions depends not only on the data supply, but also on the details of the intended use. For whom and for what purpose are the budgets and indicators made?

I began my doctoral studies fully aware that agricultural production systems alter the Earth's environment in ways that we only partially understand.

Because I find this lack of knowledge hard to accept I have felt deeply motivated learning about the knowledge gaps and trying to patch some of them. But I have also increasingly accepted that partial understanding is a condition under which life-changing decisions must be made. Inaction is also a decision. Uncertainty is not an excuse.

Humanity needs to reshape its agriculture and food systems fast, and with regard to a web of complicated and interconnected issues. I hope that this book has fulfilled the aim of contributing a little bit more certainty about some of those issues and a few ideas about how to deal with the uncertainty that remains.

References

[1] J. A. Foley, R. DeFries, G. P. Asner, et al. (2005). Global Consequences of Land Use. *Science* **309** (5734), pp. 570–574. DOI: 10.1126/science.1111772.

[2] M. Springmann, M. Clark, D. Mason-D'Croz, et al. (2018). Options for keeping the food system within environmental limits. *Nature* **562**, pp. 519–525. DOI: 10.1038/s41586-018-0594-0.

[3] L. Bouwman, K. K. Goldewijk, K. W. V. D. Hoek, A. H. W. Beusen, D. P. V. Vuuren, J. Willems, M. C. Rufino, and E. Stehfest (2013). Exploring global changes in nitrogen and phosphorus cycles in agriculture induced by livestock production over the 1900–2050 period. *Proceedings of the National Academy of Sciences* **110** (52), pp. 20882–20887. DOI: 10.1073/pnas.1012878108.

[4] L. Lassaletta, G. Billen, J. Garnier, L. Bouwman, E. Velazquez, N. D. Mueller, and J. S. Gerber (2016). Nitrogen use in the global food system: past trends and future trajectories of agronomic performance, pollution, trade, and dietary demand. *Environmental Research Letters* **11** (9), p. 095007. DOI: 10.1088/1748-9326/11/9/095007.

[5] F. Lun, J. Liu, P. Ciais, T. Nesme, J. Chang, R. Wang, D. Goll, J. Sardans, J. Peñuelas, and M. Obersteiner (2018). Global and regional phosphorus budgets in agricultural systems and their implications for phosphorus-use efficiency. *Earth System Science Data* **10** (1), pp. 1–18. DOI: 10.5194/essd-10-1-2018.

[6] E. M. Bennett, S. R. Carpenter, and N. F. Caraco (2001). Human Impact on Erodable Phosphorus and Eutrophication: A Global Perspective. *BioScience* **51** (3), pp. 227–234. DOI: 10.1641/0006-3568(2001)051[0227:HIOEPA]2.0.CO;2.

[7] J. N. Galloway, J. D. Aber, J. W. Erisman, S. P. Seitzinger, R. W. Howarth, E. B. Cowling, and B. J. Cosby (2003). The Nitrogen Cascade. *BioScience* **53** (4), pp. 341–356. DOI: 10.1641/0006-3568(2003)053[0341:TNC]2.0.CO;2.

[8] J. Elser and E. Bennett (2011). A broken biogeochemical cycle. *Nature* **478** (7367), pp. 29–31. DOI: 10.1038/478029a.

[9] M. A. Sutton, A. Bleeker, C. Howard, et al. (2013). *Our nutrient world: the challenge to produce more food and energy with less pollution.* Edinburgh: NERC/Centre for Ecology & Hydrology. 114 pp. ISBN: 978-1-906698-40-9.

[10] D. Fowler, C. E. Steadman, D. Stevenson, et al. (2015). Effects of global change during the 21st century on the nitrogen cycle. *Atmos. Chem. Phys.* **15** (24), pp. 13849–13893. DOI: 10.5194/acp-15-13849-2015.

[11] IAASTD (2009). *Synthesis report with executive summary: a synthesis of the global and sub-global IAASTD.* Ed. by B. D. McIntyre, H. R. Herren, J. Wakhungu, and R. T. Watson. Washington, D.C.: International assessment of agricultural knowledge, science and technology for development (IAASTD). ISBN: 978-1-59726-550-8.

[12] UN General Assembly (2015). *Transforming our world: the 2030 Agenda for Sustainable Development.* A/RES/70/1.

[13] FAOSTAT (2020). *New Food Balances.* URL: http://www.fao.org/faostat/en/#data/FBS (Retrieved: 2020-08-17).

[14] W. de Vries, J. Kros, C. Kroeze, and S. P. Seitzinger (2013). Assessing planetary and regional nitrogen boundaries related to food security and adverse environmental impacts. *Current Opinion in Environmental Sustainability.* Open Issue **5** (3–4), pp. 392–402. DOI: 10.1016/j.cosust.2013.07.004.

[15] S. R. Carpenter and E. M. Bennett (2011). Reconsideration of the planetary boundary for phosphorus. *Environmental Research Letters* **6** (1), p. 014009. DOI: 10.1088/1748-9326/6/1/014009.

[16] O. Oenema, H. Kros, and W. de Vries (2003). Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies. *European Journal of Agronomy.* Element Balances as Sustainability Tools **20** (1–2), pp. 3–16. DOI: 10.1016/S1161-0301(03)00067-4.

[17] O. Oenema (2015). "Nitrogen Use Efficiency (NUE) - An Indicator for the Utilisation of Nitrogen in Agricultural and Food Systems". In: *Proceedings No 733.* Meeting of the International Fertiliser Society in Cambridge, UK, on 11th December 2015. International Fertiliser Society, pp. 2–31. ISBN: 978-0-85310-410-0.

[18] S. Olofsson, E. Hjelm, and H. Nilsson (2019). “Lessons for Nitrogen Use Efficiency from the National Swedish Focus On Nutrients Project”. In: *Proceedings No 839*. Meeting of the International Fertiliser Society in Cambridge, on 12th December 2019. International Fertiliser Society, pp. 2–16. ISBN: 978-0-85310-476-6.

[19] O. Oenema (2004). Governmental policies and measures regulating nitrogen and phosphorus from animal manure in European agriculture. *Journal of Animal Science* **82** (suppl_13), E196–E206. DOI: 10.2527/2004.8213_supplE196x.

[20] Eurostat (2018). Agri-environmental indicator - gross nitrogen balance. *Statistics Explained*. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/Agri-environmental_indicator_-_gross_nitrogen_balance#Key_messages (Retrieved: 2020-08-16).

[21] Eurostat (2019). *Gross nutrient balance (aei_pr_gnb)*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/aei_pr_gnb_esms.htm (Retrieved: 2019-12-16).

[22] J. J. Schröder, H. F. M. Aarts, H. F. M. ten Berge, H. van Keulen, and J. J. Neeteson (2003). An evaluation of whole-farm nitrogen balances and related indices for efficient nitrogen use. *European Journal of Agronomy*. Element Balances as Sustainability Tools **20** (1–2), pp. 33–44. DOI: 10.1016/S1161-0301(03)00070-4.

[23] O. Godinot, M. Carof, F. Vertès, and P. Leterme (2014). SyNE: An improved indicator to assess nitrogen efficiency of farming systems. *Agricultural Systems* **127**, pp. 41–52. DOI: 10.1016/j.agsy.2014.01.003.

[24] B. Hong, D. P. Swaney, M. McCrackin, A. Svanbäck, C. Humborg, B. Gustafsson, A. Yershova, and A. Pakhomau (2017). Advances in NANI and NAPI accounting for the Baltic drainage basin: spatial and temporal trends and relationships to watershed TN and TP fluxes. *Biogeochemistry* **133** (3), pp. 245–261. DOI: 10.1007/s10533-017-0330-0.

[25] J. Le Noë, G. Billen, F. Esculier, and J. Garnier (2018). Long-term socioecological trajectories of agro-food systems revealed by N and P flows in French regions from 1852 to 2014. *Agriculture, Ecosystems & Environment* **265**, pp. 132–143. DOI: 10.1016/j.agee.2018.06.006.

[26] EEA (2005). *Agriculture and environment in EU-15 - the IRENA indicator report*. Luxembourg: Office for Official Publications of the European Communities. ISBN: 92-9167-779-5.

[27] B. Grizzetti, F. Bouraoui, and A. Aloe (2007). *Spatialised European Nutrient Balance*. Luxembourg: Office for Official Publications of the European Communities. ISBN: 978-92-79-05057-2.

[28] A. Leip, B. Achermann, G. Billen, et al. (2011). "Integrating nitrogen fluxes at the European scale". In: *The European Nitrogen Assessment*. Cambridge University Press. ISBN: 978-1-107-00612-6.

[29] W. de Vries, A. Leip, G. J. Reinds, J. Kros, J. P. Lesschen, and A. F. Bouwman (2011). Comparison of land nitrogen budgets for European agriculture by various modeling approaches. *Environmental Pollution*. Assessment of Nitrogen Fluxes to Air and Water from Site Scale to Continental Scale **159** (11), pp. 3254–3268. DOI: 10.1016/j.envpol.2011.03.038.

[30] B. Hong, D. P. Swaney, C.-M. Mörth, E. Smedberg, H. Eriksson Hägg, C. Humborg, R. W. Howarth, and F. Bouraoui (2012). Evaluating regional variation of net anthropogenic nitrogen and phosphorus inputs (NANI/NAPI), major drivers, nutrient retention pattern and management implications in the multinational areas of Baltic Sea basin. *Ecological Modelling* **227**, pp. 117–135. DOI: 10.1016/j.ecolmodel.2011.12.002.

[31] J. Le Noë, G. Billen, and J. Garnier (2017). How the structure of agro-food systems shapes nitrogen, phosphorus, and carbon fluxes: The generalized representation of agro-food system applied at the regional scale in France. *The Science of the Total Environment* **586**, pp. 42–55. DOI: 10.1016/j.scitotenv.2017.02.040.

[32] M. A. Bleken, H. Steinshamn, and S. Hansen (2005). High Nitrogen Costs of Dairy Production in Europe: Worsened by Intensification. *AMBIO: A Journal of the Human Environment* **34** (8), pp. 598–606. DOI: 10.1579/0044-7447-34.8.598.

[33] A. M. Leach, J. N. Galloway, A. Bleeker, J. W. Erisman, R. Kohn, and J. Kitzes (2012). A nitrogen footprint model to help consumers understand their role in nitrogen losses to the environment. *Environmental Development* **1** (1), pp. 40–66. DOI: 10.1016/j.envdev.2011.12.005.

[34] W. Mu, C. E. van Middelaar, J. M. Bloemhof, J. Oenema, and I. J. M. de Boer (2016). Nutrient balance at chain level: a valuable approach

to benchmark nutrient losses of milk production systems. *Journal of Cleaner Production* **112**, pp. 2419–2428. DOI: 10.1016/j.jclepro.2015.09.116.

[35] L. Lassaletta, G. Billen, B. Grizzetti, J. Anglade, and J. Garnier (2014). 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. *Environmental Research Letters* **9** (10), p. 105011. DOI: 10.1088/1748-9326/9/10/105011.

[36] B. L. Bodirsky and C. Müller (2014). Robust relationship between yields and nitrogen inputs indicates three ways to reduce nitrogen pollution. *Environmental Research Letters* **9** (11), p. 111005. DOI: 10.1088/1748-9326/9/11/111005.

[37] N. D. Mueller, L. Lassaletta, B. C. Runck, G. Billen, J. Garnier, and J. S. Gerber (2017). Declining spatial efficiency of global cropland nitrogen allocation. *Global Biogeochemical Cycles* **31** (2), 2016GB005515. DOI: 10.1002/2016GB005515.

[38] J. N. Galloway, F. J. Dentener, D. G. Capone, et al. (2004). Nitrogen Cycles: Past, Present, and Future. *Biogeochemistry* **70** (2), pp. 153–226. DOI: 10.1007/s10533-004-0370-0.

[39] E. M. Bennett and M. E. Schipanski (2013). "The Phosphorus Cycle". In: *Fundamentals of Ecosystem Science*. Elsevier, pp. 159–178. ISBN: 978-0-12-088774-3. DOI: 10.1016/B978-0-08-091680-4.00008-1.

[40] I. Öborn, A. C. Edwards, E. Witter, O. Oenema, K. Ivarsson, P. J. A. Withers, S. I. Nilsson, and A. Richert Stinzing (2003). Element balances as a tool for sustainable nutrient management: a critical appraisal of their merits and limitations within an agronomic and environmental context. *European Journal of Agronomy*. Element Balances as Sustainability Tools **20** (1–2), pp. 211–225. DOI: 10.1016/S1161-0301(03)00080-7.

[41] J. W. A. Langeveld, A. Verhagen, J. J. Neeteson, H. van Keulen, J. G. Conijn, R. L. M. Schils, and J. Oenema (2007). Evaluating farm performance using agri-environmental indicators: Recent experiences for nitrogen management in The Netherlands. *Journal of Environmental Management*. Farm Management and the Environment **82** (3), pp. 363–376. DOI: 10.1016/j.jenvman.2005.11.021.

[42] D. P. Swaney, B. Hong, C. Ti, R. W. Howarth, and C. Humborg (2012). Net anthropogenic nitrogen inputs to watersheds and riverine N

export to coastal waters: a brief overview. *Current Opinion in Environmental Sustainability*. Carbon and Nitrogen Cycles **4** (2), pp. 203–211. DOI: 10.1016/j.cosust.2012.03.004.

[43] D. Cordell, J.-O. Drangert, and S. White (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change*. Traditional Peoples and Climate Change **19** (2), pp. 292–305. DOI: 10.1016/j.gloenvcha.2008.10.009.

[44] P. Walan, S. Davidsson, S. Johansson, and M. Höök (2014). Phosphate rock production and depletion: Regional disaggregated modeling and global implications. *Resources, Conservation and Recycling* **93**, pp. 178–187. DOI: 10.1016/j.resconrec.2014.10.011.

[45] D. A. Jaffe (2000). "The Nitrogen Cycle". In: *Earth System Science*. 2nd ed. Vol. 72. International Geophysics Series. Academic Press, pp. 322–342. ISBN: 0-12-379370-X.

[46] R. A. Jahnke (2000). "The Phosphorus Cycle". In: *Earth System Science*. 2nd ed. Vol. 72. International Geophysics Series. Academic Press, pp. 360–376. ISBN: 0-12-379370-X.

[47] P. M. Groffman and E. J. Rosi-Marshall (2013). "The Nitrogen Cycle". In: *Fundamentals of Ecosystem Science*. Elsevier, pp. 137–158. ISBN: 978-0-12-088774-3. DOI: 10.1016/B978-0-08-091680-4.00008-1.

[48] K. Butterbach-Bahl, P. Gundersen, P. Ambus, et al. (2011). "Nitrogen processes in terrestrial ecosystems". In: *The European Nitrogen Assessment*. Cambridge University Press. ISBN: 978-1-107-00612-6.

[49] A. McNeill and M. Unkovich (2007). "The Nitrogen Cycle in Terrestrial Ecosystems". In: *Nutrient Cycling in Terrestrial Ecosystems*. Ed. by D. P. Marschner and P. D. Z. Rengel. Soil Biology 10. Springer Berlin Heidelberg, pp. 37–64. ISBN: 978-3-540-68026-0 978-3-540-68027-7. DOI: 10.1007/978-3-540-68027-7_2.

[50] B. Thamdrup (2012). New Pathways and Processes in the Global Nitrogen Cycle. *Annual Review of Ecology, Evolution, and Systematics* **43** (1), pp. 407–428. DOI: 10.1146/annurev-ecolsys-102710-145048.

[51] R. Einarsson (2017). *Assessing reactive nitrogen flows in European agricultural systems*. Licentiate Thesis. Chalmers University of Technology. URL: https://publications.lib.chalmers.se/publication/252079.

[52] J. W. Erisman, J. N. Galloway, S. Seitzinger, A. Bleeker, N. B. Dise, A. M. R. Petrescu, A. M. Leach, and W. de Vries (2013). Consequences

of human modification of the global nitrogen cycle. *Phil. Trans. R. Soc. B* **368** (1621), p. 20130116. DOI: 10.1098/rstb.2013.0116.

[53] M. A. Sutton, C. M. Howard, J. W. Erisman, G. Billen, A. Bleeker, P. Grennfelt, H. van Grinsven, and B. Grizzetti, eds. (2011). *The European Nitrogen Assessment: Sources, Effects and Policy Perspectives.* Cambridge: Cambridge University Press. ISBN: 978-0-511-97698-8.

[54] F. E. Allison (1955). "The Enigma of Soil Nitrogen Balance Sheets". In: *Advances in Agronomy.* Ed. by A. G. Norman. Vol. 7. Academic Press, pp. 213–250. DOI: 10.1016/S0065-2113(08)60339-9.

[55] C. R. Frink (1969). Water Pollution Potential Estimated from Farm Nutrient Budgets. *Agronomy Journal* **61** (4), pp. 550–553. DOI: 10.2134/agronj1969.00021962006100040020x.

[56] W. H. Garman (1970). "Agricultural Nutrient Budget". In: *Nutrient Mobility in Soils: Accumulation and Losses.* John Wiley & Sons, Ltd, pp. 61–74. ISBN: 978-0-89118-899-5. DOI: 10.2136/sssaspecpub4.c4.

[57] N. van Breemen, P. A. Burrough, E. J. Velthorst, H. F. van Dobben, T. de Wit, T. B. Ridder, and H. F. R. Reijnders (1982). Soil acidification from atmospheric ammonium sulphate in forest canopy throughfall. *Nature* **299** (5883) (5883), pp. 548–550. DOI: 10.1038/299548a0.

[58] H. F. M. Aarts, E. E. Biewing, and H. van Keulen (1992). Dairy farming systems based on efficient nutrient management. *Netherlands Journal of Agricultural Science* **40** (3) (3), pp. 285–299. DOI: 10.18174/njas.v40i3.16514.

[59] T. Dalgaard, J. F. Bienkowski, A. Bleeker, et al. (2012). Farm nitrogen balances in six European landscapes as an indicator for nitrogen losses and basis for improved management. *Biogeosciences* **9** (12), pp. 5303–5321. DOI: 10.5194/bg-9-5303-2012.

[60] Eurostat (2013). *Nutrient Budgets – Methodology and Handbook. Version 1.02.* Luxembourg: Eurostat and OECD.

[61] A. Leip, W. Britz, F. Weiss, and W. de Vries (2011). Farm, land, and soil nitrogen budgets for agriculture in Europe calculated with CAPRI. *Environmental Pollution.* Assessment of Nitrogen Fluxes to Air and Water from Site Scale to Continental Scale **159** (11), pp. 3243–3253. DOI: 10.1016/j.envpol.2011.01.040.

[62] R. W. Howarth, G. Billen, D. Swaney, et al. (1996). Regional nitrogen budgets and riverine N & P fluxes for the drainages to the North Atlantic Ocean: Natural and human influences. *Biogeochemistry.* Ed.

by R. W. Howarth, pp. 75–139. DOI: 10.1007/978-94-009-1776-7_3.

[63] M. Unkovich (2008). *Measuring Plant-associated Nitrogen Fixation in Agricultural Systems*. Australian Centre for International Agricultural Research. 258 pp. ISBN: 978-1-921531-26-2. Google Books: rD8TtwAACAAJ.

[64] D. F. Herridge, M. B. Peoples, and R. M. Boddey (2008). Global inputs of biological nitrogen fixation in agricultural systems. *Plant and Soil* **311** (1-2), pp. 1–18. DOI: 10.1007/s11104-008-9668-3.

[65] Eurostat (2019). *Crop production (apro_cp)*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/apro_cp_esms.htm (Retrieved: 2020-01-15).

[66] H. Høgh-Jensen, R. Loges, F. V. Jørgensen, F. P. Vinther, and E. S. Jensen (2004). An empirical model for quantification of symbiotic nitrogen fixation in grass-clover mixtures. *Agricultural Systems* **82** (2), pp. 181–194. DOI: 10.1016/j.agsy.2003.12.003.

[67] J. Anglade, G. Billen, and J. Garnier (2015). Relationships for estimating N2 fixation in legumes: incidence for N balance of legume-based cropping systems in Europe. *Ecosphere* **6** (3), pp. 1–24. DOI: 10.1890/ES14-00353.1.

[68] A. Mulier, G. Hofman, E. Baecke, et al. (2003). A methodology for the calculation of farm level nitrogen and phosphorus balances in Flemish agriculture. *European Journal of Agronomy*. Element Balances as Sustainability Tools **20** (1), pp. 45–51. DOI: 10.1016/S1161-0301(03)00071-6.

[69] J. Kros, G. B. M. Heuvelink, G. J. Reinds, J. P. Lesschen, V. Ioannidi, and W. De Vries (2012). Uncertainties in model predictions of nitrogen fluxes from agro-ecosystems in Europe. *Biogeosciences* **9** (11), pp. 4573–4588. DOI: 10.5194/bg-9-4573-2012.

[70] J. Oenema, S. Burgers, H. van Keulen, and M. van Ittersum (2015). Stochastic uncertainty and sensitivities of nitrogen flows on dairy farms in The Netherlands. *Agricultural Systems* **137**, pp. 126–138. DOI: 10.1016/j.agsy.2015.04.009.

[71] G. K. MacDonald and E. M. Bennett (2009). Phosphorus Accumulation in Saint Lawrence River Watershed Soils: A Century-Long Perspective. *Ecosystems* **12** (4), pp. 621–635. DOI: 10.1007/s10021-009-9246-4.

[72] K. C. van Dijk, J. P. Lesschen, and O. Oenema (2016). Phosphorus flows and balances of the European Union Member States. *Science of The Total Environment.* Special Issue on Sustainable Phosphorus Taking Stock: Phosphorus Supply from Natural and Anthropogenic Pools in the 21st Century **542** (Part B), pp. 1078–1093. DOI: 10.1016/j.scitotenv.2015.08.048.

[73] S. M. Powers, T. W. Bruulsema, T. P. Burt, et al. (2016). Long-term accumulation and transport of anthropogenic phosphorus in three river basins. *Nature Geoscience* **9** (5) (5), pp. 353–356. DOI: 10.1038/ngeo2693.

[74] J. Némery and J. Garnier (2016). The fate of phosphorus. *Nature Geoscience* **9** (5) (5), pp. 343–344. DOI: 10.1038/ngeo2702.

[75] G. Tóth, R.-A. Guicharnaud, B. Tóth, and T. Hermann (2014). Phosphorus levels in croplands of the European Union with implications for P fertilizer use. *European Journal of Agronomy* **55**, pp. 42–52. DOI: 10.1016/j.eja.2013.12.008.

[76] L. Bergström, H. Kirchmann, F. Djodjic, et al. (2015). Turnover and Losses of Phosphorus in Swedish Agricultural Soils: Long-Term Changes, Leaching Trends, and Mitigation Measures. *Journal of Environmental Quality* **44** (2), pp. 512–523. DOI: 10.2134/jeq2014.04.0165.

[77] A. Svanbäck, B. Ulén, A. Etana, L. Bergström, P. J. A. Kleinman, and L. Mattsson (2013). Influence of soil phosphorus and manure on phosphorus leaching in Swedish topsoils. *Nutrient Cycling in Agroecosystems* **96** (2), pp. 133–147. DOI: 10.1007/s10705-013-9582-9.

[78] F. Worrall, N. J. K. Howden, and T. P. Burt (2015). Evidence for nitrogen accumulation: the total nitrogen budget of the terrestrial biosphere of a lowland agricultural catchment. *Biogeochemistry* **123** (3), pp. 411–428. DOI: 10.1007/s10533-015-0074-7.

[79] T. A. Clair, N. Pelletier, S. Bittman, et al. (2014). Interactions between reactive nitrogen and the Canadian landscape: A budget approach. *Global Biogeochemical Cycles* **28** (11), 2014GB004880. DOI: 10.1002/2014GB004880.

[80] Science Advisory Board (2011). *Reactive Nitrogen in the United States: An Analysis of Inputs, Flows, Consequences and Management Options.* United States Environmental Protection Agency, Science Advisory Board. 140 pp. URL: https://yosemite.epa.gov/sab/sabproduct.nsf/WebBOARD/INCFullReport/$File/Final%20INC%

20Report_8_19_11(without%20signatures).pdf (Retrieved: 2017-06-05).

[81] K. J. Van Meter, N. B. Basu, J. J. Veenstra, and C. L. Burras (2016). The nitrogen legacy: emerging evidence of nitrogen accumulation in anthropogenic landscapes. *Environmental Research Letters* **11** (3), p. 035014. DOI: 10.1088/1748-9326/11/3/035014.

[82] V. Smil (1999). Nitrogen in crop production: An account of global flows. *Global Biogeochemical Cycles* **13** (2), pp. 647–662. DOI: 10.1029/1999GB900015.

[83] G. Mitchell, A. May, and A. McDonald (1995). PICABUE: a methodological framework for the development of indicators of sustainable development. *International Journal of Sustainable Development & World Ecology* **2** (2), pp. 104–123. DOI: 10.1080/13504509509469893.

[84] H. M. G. van der Werf and J. Petit (2002). Evaluation of the environmental impact of agriculture at the farm level: a comparison and analysis of 12 indicator-based methods. *Agriculture, Ecosystems & Environment* **93** (1), pp. 131–145. DOI: 10.1016/S0167-8809(01)00354-1.

[85] European Commission (2006). *COM (2006) 508 final.* URL: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:52006DC0508 (Retrieved: 2019-12-16).

[86] K. Butterbach-Bahl, E. M. Baggs, M. Dannenmann, R. Kiese, and S. Zechmeister-Boltenstern (2013). Nitrous oxide emissions from soils: how well do we understand the processes and their controls? *Phil. Trans. R. Soc. B* **368** (1621), p. 20130122. DOI: 10.1098/rstb.2013.0122.

[87] G. Billen, M. Silvestre, B. Grizzetti, et al. (2011). "Nitrogen flows from regional European watersheds to coastal marine waters". In: *The European Nitrogen Assessment.* Cambridge University Press, pp. 271–297. ISBN: 978-1-107-00612-6.

[88] B. Hansen, T. Dalgaard, L. Thorling, B. Sørensen, and M. Erlandsen (2012). Regional analysis of groundwater nitrate concentrations and trends in Denmark in regard to agricultural influence. *Biogeosciences* **9** (8), pp. 3277–3286. DOI: 10.5194/bg-9-3277-2012.

[89] W. de Vries, P. Cellier, J. W. Erisman, and M. A. Sutton (2011). Assessment of nitrogen fluxes to air and water from site scale to continental scale: An overview. *Environmental Pollution.* Assessment of Nitrogen

Fluxes to Air and Water from Site Scale to Continental Scale **159** (11), pp. 3143–3148. DOI: 10.1016/j.envpol.2011.08.047.

[90] K. Parris (1998). Agricultural nutrient balances as agri-environmental indicators: an OECD perspective. *Environmental Pollution*. Nitrogen, the Confer-N-s First International Nitrogen Conference 1998 **102** (1, Supplement 1), pp. 219–225. DOI: 10.1016/S0269-7491(98)80036-5.

[91] A. Leip, F. Weiss, J. P. Lesschen, and H. Westhoek (2014). The nitrogen footprint of food products in the European Union. *The Journal of Agricultural Science* **152**, S20–S33. DOI: 10.1017/S0021859613000786.

[92] H. Shibata, J. N. Galloway, A. M. Leach, et al. (2017). Nitrogen footprints: Regional realities and options to reduce nitrogen loss to the environment. *Ambio* **46** (2), pp. 129–142. DOI: 10.1007/s13280-016-0815-4.

[93] D. M. Bailer-Jones (2002). Scientists' Thoughts on Scientific Models. *Perspectives on Science* **10** (3), pp. 275–301. DOI: 10.1162/106361402321899069.

[94] R. Frigg and S. Hartmann (2020). "Models in Science". In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Spring 2020. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/spr2020/entries/models-science/ (Retrieved: 2020-08-26).

[95] Eurostat (2019). *Annual Crop Statistics Handbook*. 2019 Edition. Eurostat. 166 pp. URL: https://ec.europa.eu/eurostat/cache/metadata/Annexes/apro_cp_esms_an1.pdf (Retrieved: 2019-05-16).

[96] Eurostat (2017). *Farm structure (ef)*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/ef_esms.htm (Retrieved: 2020-01-15).

[97] Eurostat (2019). *Sales of manufactured fertilizers (source: Fertilizers Europe) (aei_fm_manfert)*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/aei_fm_manfert_esms.htm (Retrieved: 2020-01-09).

[98] Eurostat (2019). *Consumption of inorganic fertilizers (aei_fm_usefert)*. URL: https://ec.europa.eu/eurostat/cache/metadata/en/aei_fm_usefert_esms.htm (Retrieved: 2020-01-09).

[99] Eurostat (2012). *NUTS 2010 GIS data*. URL: https://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units/nuts (Retrieved: 2020-01-15).

[100] M. Bossard, J. Feranec, and J. Otahel (2000). *CORINE land cover technical guide: Addendum 2000. Report 40*. European Environment Agency. URL: http://www.eea.europa.eu/publications/tech40add (Retrieved: 2015-09-04).

[101] T. P. Robinson, G. R. W. Wint, G. Conchedda, T. P. Van Boeckel, V. Ercoli, E. Palamara, G. Cinardi, L. D'Aietti, S. I. Hay, and M. Gilbert (2014). Mapping the Global Distribution of Livestock. *PLoS ONE* **9** (5), e96084. DOI: 10.1371/journal.pone.0096084.

[102] UNFCCC (2019). *National Inventory Submissions 2019*. URL: https://unfccc.int/process-and-meetings/transparency-and-reporting/reporting-and-review-under-the-convention/greenhouse-gas-inventories-annex-i-parties/national-inventory-submissions-2019 (Retrieved: 2020-06-24).

[103] FADN (2019). *FADN Standard Results Database, May 2019 version*. URL: https://ec.europa.eu/agriculture/rica/database/reports/archives/fadn20190531.zip (Retrieved: 2020-01-09).

[104] FADN (2015). *FADN GIS data*. URL: https://ec.europa.eu/agriculture/rica/images/othermaps/FADN_RICA_PL_2012_20M_shape.7z (Retrieved: 2020-01-15).

[105] G. L. Velthof (2014). *Task 1 of Methodological studies in the field of Agro-Environmental Indicators. Lot 1 excretion factors. Final draft*. Wageningen: Alterra. URL: http://ec.europa.eu/eurostat/documents/2393397/8259002/LiveDate_2014_Task1.pdf/e1ac8f30-3c76-4a61-b607-de99f98fc7cd (Retrieved: 2020-01-09).

[106] INRAE, CIRAD, AFZ, and FAO (2020). *List of feeds*. URL: https://www.feedipedia.org/content/feeds (Retrieved: 2020-08-27).

[107] FAOSTAT (2020). *FAO Statistics (FAOSTAT)*. URL: http://www.fao.org/faostat/en/#home (Retrieved: 2020-05-15).

[108] FAOSTAT (2020). *Definitions and standards for FAOSTAT crop production data*. URL: http://www.fao.org/faostat/en/#data/QC (Retrieved: 2020-06-25).

[109] FAOSTAT (2016). *Metadata on crop statistics*. URL: http://www.fao.org/faostat/en/#data/QC/metadata (Retrieved: 2020-06-16).

[110] FAOSTAT (2016). *Fertilizers Archive Metadata*. URL: http://www.fao.org/faostat/en/#data/RA (Retrieved: 2020-06-15).

[111] FAOSTAT (2019). *Fertilizers by Nutrient Methodology*. URL: http://fenixservices.fao.org/faostat/static/documents/RFN/RFN_EN_README.pdf (Retrieved: 2020-06-15).

[112] FAOSTAT (2020). *Definitions and standards for FAOSTAT land use data*. URL: http://www.fao.org/faostat/en/#data/RL (Retrieved: 2020-06-25).

[113] IFASTAT (2020). *IFASTAT Fertilizer Consumption Database*. URL: https://www.ifastat.org/faq (Retrieved: 2020-06-15).

[114] O. Häggström (2007). *Uniform distribution is a model assumption*. URL: http://www.math.chalmers.se/~olleh/reply_to_Dembski.pdf (Retrieved: 2020-08-26).

[115] V. Verendel and O. Häggström (2017). Fermi's paradox, extraterrestrial life and the future of humanity: a Bayesian analysis. *International Journal of Astrobiology* **16** (1), pp. 14–18. DOI: 10.1017/S1473550415000452.

[116] R. D. Peng (2011). Reproducible Research in Computational Science. *Science* **334** (6060), pp. 1226–1227. DOI: 10.1126/science.1213847.

[117] J. M. Perkel (2020). Challenge to scientists: does your ten-year-old code still run? *Nature* **584** (7822) (7822), pp. 656–658. DOI: 10.1038/d41586-020-02462-7.

[118] R. Einarsson (2017). *Source code for Estimation of the biogas production potental from manure and crop residues in the EU*. URL: https://github.com/rasmuse/biogas-residues-manure (Retrieved: 2020-08-26).

[119] R. Einarsson (2020). *Eurostat tools (eust) v0.5.1*. URL: https://github.com/rasmuse/eust (Retrieved: 2020-08-26).

[120] R. Einarsson (2020). *Source code for subnational P budgets in EU28 for 2013*. Zenodo. DOI: 10.5281/zenodo.3610359.

[121] G. Van Rossum and F. L. Drake (2009). *Python 3 reference manual*. Scotts Valley, CA: CreateSpace. ISBN: 1-4414-1269-7.

[122] R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. URL: https://www.R-project.org/.

[123] P. Virtanen, R. Gommers, T. E. Oliphant, et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* **17** (3) (3), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[124] S. van der Walt, S. C. Colbert, and G. Varoquaux (2011). The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science Engineering* **13** (2), pp. 22–30. DOI: 10.1109/MCSE.2011.37.

[125] W. McKinney (2010). Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*, pp. 56–61. DOI: 10.25080/Majora-92bf1922-00a.

[126] J. D. Hunter (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* **9** (3), pp. 90–95. DOI: 10.1109/MCSE.2007.55.

[127] F. Pedregosa, G. Varoquaux, A. Gramfort, et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **12** (85), pp. 2825–2830. URL: http://jmlr.org/papers/v12/pedregosa11a.html (Retrieved: 2020-08-26).

[128] The Git community (n.d.). *Git*. URL: https://git-scm.com/.

[129] Corporation for Digital Scholarship (n.d.). *Zotero*. URL: https://www.zotero.org/.

[130] W. Anderson, L. You, S. Wood, U. Wood-Sichra, and W. Wu (2015). An analysis of methodological and spatial differences in global cropping systems models and maps. *Global Ecology and Biogeography* **24** (2), pp. 180–191. DOI: 10.1111/geb.12243.

[131] H. J. Smit, M. J. Metzger, and F. Ewert (2008). Spatial distribution of grassland productivity and land use in Europe. *Agricultural Systems* **98** (3), pp. 208–219. DOI: 10.1016/j.agsy.2008.07.004.

[132] S. G. Sommer, J. Webb, and N. D. Hutchings (2019). New Emission Factors for Calculation of Ammonia Volatilization From European Livestock Manure Management Systems. *Frontiers in Sustainable Food Systems* **3**. DOI: 10.3389/fsufs.2019.00101.

[133] O. Gavrilova, A. Leip, H. Dong, et al. (2019). "Emissions from livestock and manure management". In: *2019 Refinement to the 2006*

IPCC Guidelines for National Greenhouse Gas Inventories. Switzerland: IPCC. ISBN: 978-4-88788-232-4.

[134] H. Menzi, T. Kupper, and E. Spiess (2015). "Example of a country manure management profile – Switzerland". In: *Rural-Urban Symbiosis*. RAMIRAN 2015 – 16th International Conference. Hamburg, Germany, p. 12. ISBN: 978-3-941492-95-0.

[135] EMEP Centre on Emission Inventories and Projections (2020). *Status of reporting to the LRTAP convention. 2019 data*. URL: `https://www.ceip.at/status-of-reporting-and-review-results/2019-submissions` (Retrieved: 2020-08-28).

[136] EMEP Centre on Emission Inventories and Projections (2020). *Officially reported activity data*. URL: `https://www.ceip.at/webdab-emission-database/officially-reported-activity-data` (Retrieved: 2020-08-28).

[137] Á. Patay and K. Lovas (2017). *Pilot studies to develop methodological improvements to agri- environmental statistics*. URL: `http://ec.europa.eu/eurostat/documents/2393397/8259002/Grant_2016_HU_Final+report.pdf/7afe66f7-5049-421b-9159-afb692c10df8` (Retrieved: 2020-08-28).

[138] J. Liu, H. Mooney, V. Hull, et al. (2015). Systems integration for global sustainability. *Science* **347** (6225). DOI: `10.1126/science.1258832`.

[139] F. Pendrill, M. Persson, J. Godar, T. Kastner, D. Moran, S. Schmidt, and R. Wood (2019). Agricultural and forestry trade drives large share of tropical deforestation emissions. *Global Environmental Change* **56**, pp. 1–10. DOI: `10.1016/j.gloenvcha.2019.03.002`.