Reviewer 2

Overall comment: This study developed a dataset of annual P surplus across Europe at a spatial resolution of 5 arcmin during 1850-2019. The uncertainties of P surplus estimation were considered by using two fertilizer estimates, six animal manure estimates, and two cropland and two pasture P removal estimates. Country-level survey data and multiple spatial maps were used to develop this dataset. The manuscript provided a very detailed description of the methodology and was easily understood. However, I still have several concerns regarding the novelty of the dataset and the reliability of the methods.

Reply: We thank the reviewer for their careful review of the manuscript and their useful comments. We reply below to the points raised by the Reviewer.

1.1 — The published 48 P surplus estimates were not very useful, since most users may only use the ensemble mean. I suggest authors publish data of all P input and output variables instead of only publishing P surplus data.

Reply:

Thank you for your insightful comments regarding the publication of the 48 P surplus estimates. We acknowledge the reviewer's concern about the potential over-reliance on ensemble mean estimates by some users and the suggestion to publish additional datasets. In response, we now provide not only the 48 P surplus estimates but also detailed datasets on P inputs and P outputs, which will allow users to carry out more targeted analyses based on specific input and output variables.

The objective of our study is to focus on long-term P surplus dynamics across Europe. Therefore, publishing the P surplus estimates remains crucial for understanding the broader environmental implications. However, we agree that providing access to the underlying input and output data significantly enhances the dataset's utility for the broader research community. **Importantly, the inclusion of all 48 estimates is a key strength of our study, as it accounts for uncertainties inherent in P surplus estimation—an aspect often overlooked in the majority of previous studies.** The uncertainty in P surplus estimates arises from various factors, including different assumptions about fertilizer and manure distribution to cropland and pasture, crop uptake coefficients, and historical data quality. The importance of accounting for uncertainty is increasingly emphasized in nutrient research, as demonstrated by recent work, such as Guejjoud et al. (2023); Ringeval et al. (2024), Sarrazin et al. (2024) and Zhang et al. (2021), which underscores how datasets incorporating uncertainty provide a stronger foundation for robust assessment of underlying nutrient dynamics. Indeed, in a recent comprehensive analysis of nitrogen budget compilation, Zhang et al. (2021) emphasizes that *"To improve nitrogen budget estimates, current uncertainties in concepts, data, and methods need to be addressed . . . "*. While the study focused on the nitrogen budget, similar recommendations can be applied to phosphorus budget estimates, as the main underlying components such as crop removal, mineral fertilizer and manure, are consistent among both nutrients. Our study contributes to this direction and takes a step toward uncertainty assessment in the P surplus budget.

While it may happen that many users may rely on the ensemble mean, we believe that providing uncertainty estimates is a good scientific practice, as it offers a more comprehensive understanding of reconstructed datasets. This is true especially in our case for reconstructing the past nutrient budgets where many methodological aspects and underlying components remain uncertain (e.g., fertilizer or manure applications on different land-use types) (Zhang et al., 2021). To this end, by including the uncertainty estimates alongside the ensemble mean P surplus estimates allows users to assess the range of variability. Although we cannot control how users choose to utilize the data, we have provided recommendations and made all the datasets available to facilitate informed selections/choices.

To reflected these points, we have further emphasized about the importance of accounting for uncertainty in the introduction of the revised manuscript in line 51-60.

Nutrient budgets tend to have large uncertainties (Zhang et al., 2021; Ludemann et al., 2023). Uncertainties in P budgets can stem from limited knowledge about the distribution of mineral fertilizers and animal manure on cropland and pasture and about the P removal coefficients, among other factors (Ludemann et al., 2023). As a result, the different studies of Table 2 and Table 3 adopted different schemes to allocate mineral fertilizer and animal manure to cropland and different coefficient values. While some studies explicitly consider uncertainties (e.g., Guejjoud et al. (2023); Antikainen et al. (2008); Lun et al. (2018); Muntwyler et al. (2024); Ringeval et al. (2024); Ludemann et al. (2023); Panagos et al. (2022), listed in Tables 2 and 3), the majority do not. Ignoring this uncertainty could lead to inaccurate assessments of P dynamics and, consequently, flawed policy recommendations (Oenema et al., 2003). Recent studies, such as Guejjoud et al. (2023); Ringeval et al. (2024), Sarrazin et al. (2024) and Zhang et al. (2021) underscore the need for uncertainty-aware nutrient datasets to support quantification of nutrient budgets and robust water quality assessments.

In addition to publishing the data for P inputs and P outputs, we have also enhanced the manuscript with new visualizations. Figure R1 (which corresponds to Figure 2 in the revised manuscript) now includes gridded maps of P inputs and outputs (alongside P surplus), offering a more detailed spatial overview. We have also updated Figure R2 (which corresponds to Figure 5 in the revised manuscript), which now illustrates the decadal trajectories of agricultural and total P inputs (shown in orange) and P outputs (in blue), while the P surplus is represented by a red line for each decade. Figure R3 and Figure R4 (which corresponds to Supplementary Figure S8 and S9, respectively) further shows decadal trajectories of agricultural and total P inputs for different European countries. This provides a clearer illustration of how inputs and outputs have evolved over time and contributed to changing P surplus levels.

In summary, while the focus of our study is on P surplus, the additional datasets of P inputs and outputs, along with the ensemble of 48 estimates, offer a comprehensive tool for researchers interested in exploring P dynamics in greater depth. We believe that this expanded data access addresses the reviewer's concerns and enhances the dataset's utility, offering researchers a broader set of resources for studying P dynamics in greater depth. Thank you for your constructive feedback on this aspect.

1.2 — The authors claimed that the novelty of this data is considering P surplus on non-agricultural land. First, it is very weird to identify P inputs (atmospheric deposition and weathering) on non-agricultural land as "surplus". Second, there are no figures showing the results of P surplus on non-agricultural land. The inputs of deposition and weathering puts of P are very low compared to fertilizer and manure inputs, and that is one of the reasons the results in this study are very close to Zou et al. and Ludemann et al. Third, the calculation of P weathering in urban is uncommon. Hartmann's data was developed on soil, and it cannot be directly used on impervious land. Fourth, aside from forests, semi-natural vegetation, and urban areas, what about shrubland and other land use types? Overall, calculating P surplus on non-agricultural land does not make this dataset distinct from other previous datasets.

Reply:

Thank you for your valuable feedback and for raising important points regarding the calculation and importance of P surplus budget over non-agricultural areas. We have carefully considered each aspect you highlighted, and below, we provide detailed responses and clarifications for each of your points.

- **Terminology for P surplus on non-agricultural land:** We have chosen to use the term **P surplus** across all land types (both agricultural and non-agricultural) to ensure consistency throughout the manuscript. This unified terminology aligns with our broader aim of analyzing phosphorus dynamics across all diverse land use types. It highlights the importance of understanding both inputs and outputs in various land types, including in areas where agriculture P sources are not dominant. P inputs from other sources may still contribute to long-term environmental challenges. Furthermore, studies on water quality assessment (e.g. Van Meter et al., 2021), requires P surplus data not just from agricultural areas but also from other areas/sources to quantify and analyze total catchment P export. We acknowledge the reviewer's concern regarding the potential confusion in referring to P inputs as P surplus over non-agricultural areas. To address this, we have provided a clarification where this term first appears in the revised manuscript at line 63 in Section 1.
- **Non-agricultural P surplus contributions:** We would like to clarify and explain the relevance of providing P surplus on non-agricultural soils. Our analysis shows that the **northern European countries** (Norway, Sweden, Finland) have a higher contribution of non-agricultural P surplus, mainly over the forests and semi-natural vegetation, compared to other regions. In Norway, non-agricultural areas accounted for approximately 28–40% of the total P surplus in the 1850s, and though this decreased to around 13% in recent decades, it remains notable compared to other regions. Similarly, Finland exhibited a non-agricultural P surplus contribution of 36–42% in the 1850s, declining to 19–31% by 2000–2019. We also have added Figures R5 (Figure 3) and R6-R7 (Supplementary Figures S6 and S7) to illustrate this contribution. Although the non-agricultural P surplus is modest compared to the agricultural P surplus in these countries, it is nevertheless important for understanding the overall P budget, especially in regions with substantial forested and semi-natural landscapes. By accounting for different sources of P inputs and outputs across all considered landscapes, our dataset can support further environmental assessments, such as analyzing the fate of P surplus in catchment-wide water quality studies. Other sources of P inputs to river systems, such as point sources (e.g., wastewater treatment plants), are not considered here. However, our group has worked on reconstructing point source nutrient inputs in a separate study (see (Sarrazin et al., 2024)).
- **Comparison with previous studies**: We also would like to clarify that the datasets of Zou et al. (2022) and Ludemann et al. (2023) mentioned by the Reviewer provide P surplus for **cropland** only (as reported for instance in Table 2, Section 3.2 and in the captions of Figures 6 and 7 of our manuscript). Therefore, the comparison between the datasets of Zou et al. (2022) and Ludemann et al. (2023) and our dataset is limited to croplands and cannot be extended to the relevance of P surplus budgets for other areas, including pasture-dominated lands within agricultural regions and non-agricultural areas.
- **P weathering in urban areas:** We recognize the uncommon application of Hartmann's soil-based weathering data to impervious urban areas. In response to your comment, we have excluded P weathering from urban areas in our

revised estimates. While urban P surplus remains a relatively minor contributor, this adjustment ensures that our calculations align more closely with realistic land use characteristics.

- **Clarification on shrubland and other land types:** In response to your query on shrublands and other land types, we clarify that **semi-natural vegetation in our dataset includes shrubland**. Specifically, for non-agricultural areas in our dataset, we used the classification of land cover categories from global land cover (GLC) (Bartholomé and Belward, 2005) that is available at a spatial resolution of 300 m and includes 23 land cover classes. From these classes, we selected all relevant land types representing non-agricultural land, which we divided into four categories, namely semi-natural-vegetation (tree, shrub-land, herbaceous cover), forest (broad-leaved, evergreen and deciduous forest), non-vegetation (bare areas, water bodies) and urban areas. This ensures that all non-agricultural areas contributing to P surplus are represented in our analysis.

- **Distinction from previous datasets:** Our dataset offers several unique contributions that collectively provide a more comprehensive view of the phosphorus (P) surplus budget across Europe. While including non-agricultural land types is crucial for gaining a thorough understanding of the total P surplus budget within a given landscape unit (e.g., catchments or administrative units), our approach is further enhanced by additional aspects such as extended spatial and temporal coverage and the incorporation of uncertainty assessments in the P surplus budgets. Below, we outline some key features that distinguish our dataset from existing ones or complement them beyond considering both agricultural and non-agricultural areas:

  - As highlighted in our first reply to the reviewer, our dataset includes 48 distinct P surplus estimates, providing a comprehensive representation of uncertainty—a critical aspect partially considered in some past studies reported in Tables 2 and 3. This uncertainty-aware approach is essential for reliably assessing the evolutions of P dynamics and supporting effective policy recommendations, as highlighted in recent literature on nutrient budgeting (Guejjoud et al., 2023; Sarrazin et al., 2024; Zhang et al., 2021).
  - With a spatio-temporal coverage at gridded scale and extending from 1850 to 2019, our dataset provides a long-term, detailed view of P surplus budget across diverse European landscapes. Further, in comparison to existing global databases, such as those by Zou et al. (2022) and Ludemann et al. (2023), which are limited to cropland P budgets, our dataset extends further by covering both cropland and pastureland P budgets. This provides a more comprehensive view of P dynamics across agricultural areas, while covering P budgets over the non-agricultural areas as well. The spatial and temporal depth enhance the utility of our dataset for several studies, ranging from historical agriculture and environmental change to biogeochemical cycling, nutrient management, legacy store characterizations, and water quality assessments, where understanding shifts over time is critical.

In response to your feedback, we have incorporated new gridded and country-level visualizations to better emphasis the contribution of non-agricultural P surplus to total P surplus.

Figure R5 (Figure 3 in the revised manuscript) displays this contribution at the gridded level, while Figures R6 and R7 (Supplementary Figures S6 and S7) depict contributions by countries and over time. To further elaborate on this point, we have added the following text in section 3.1 (at line 651) of the revised manuscript:

The importance of non-agricultural P surplus is highlighted in Fig R5, which illustrates its contribution to total P surplus. Northern European countries, such as Norway, Sweden, and Finland, show a higher contribution of non-agricultural P surplus, with 30–60% contribution across 70% of grid cells during the entire period (1850–2019). Central and Western Europe exhibit more variable contributions over time. For example, in 1900 and 1930, the non-agricultural contribution in these regions ranged between 10–30%, but it decreased to around 10% by 1990, with further declines in recent years. Southern Europe, meanwhile, displayed a moderate and stable contribution of up to 20% from 1960 to 2019. Supplementary Figures S6 and S7 provide additional insights, showing the contribution of non-agricultural P surplus *at both the country level and on a decadal scale. Northern and Eastern European countries demonstrate increasing contributions over time, such as Estonia (from 15% in 1850–60 to 30% in 2010–19) and Sweden (from 35% to 40% over the same period). Meanwhile, countries like Belgium, the Netherlands, and Switzerland show a consistent decrease in contribution throughout the period, such as Switzerland dropping from 40% in 1850–60 to 5% in 2010–19.Understanding these dynamics is critical for devising holistic nutrient management strategies that account for the role of non-agricultural P sources. By incorporating non-agricultural P surplus data, our dataset enables a more comprehensive understanding of P fluxes across Europe.*

To clarify the shrubland and other land use category in our dataset and to further explain our approach to constructing non-agricultural land, we have revised the text in section 2.2.2 at line 206 of the revised manuscript, which now reads as follows:

The non-agricultural area in a grid cell was calculated as the remaining area after allocating cropland and pasture areas. We used the classification of land cover categories from global land cover (GLC) (Bartholomé and Belward, 2005) that is available at a spatial resolution of 300 m. GLC includes 23 land cover classes that we grouped into 5 categories namely, cropland, semi-natural-vegetation (i.e. vegetation not planted by humans but influenced by human actions (Di Gregorio, 2005) including tree, shrub-land, herbaceous cover, Lichen and mosses), forest (broad-leaved, evergreen and deciduous forest), non-vegetation (bare areas, water bodies) and urban area. The proportions of these categories were then applied to the non-agricultural area to estimate their annual development from 1850 to 2019.

Moreover, to emphasize the unique aspects of our dataset, we have added a detailed section 4 in the revised manuscript, in which we highlight the value of our dataset at line 875-923.

1.3 — The fertilizer data before 1960 was calculated by using the temporal changes from Holland et al. However, Holland only provides N fertilizer data. The N fertilizer is produced from the Haber-Bosch process while P fertilizer is produced from mineral rock. These two different technology may not lead to a constant N:P ratio of fertilizer before 1960. Therefore, it is not a solid method to directly use temporal changes of N fertilizer on P fertilizer.

Reply:

Thank you for your valuable feedback. We fully acknowledge the distinction between the production processes of nitrogen (N) fertilizers and phosphorus (P) fertilizers, as N fertilizers are derived from the Haber-Bosch process while P fertilizers are produced from phosphate rock. Given these technological differences, we recognize that it is not appropriate to assume a constant N:P ratio for fertilizers before 1960.

In response to your comment, we have revised our methodology. Instead of relying on N fertilizer trends from Holland et al. (2005) as a proxy for changes in P fertilizer, we have now incorporated a **global dataset (Cordell et al., 2009) that traces the historical sources of phosphorus fertilizers from phosphate rock (1800–2000)**. This dataset provides a more reliable temporal trend for P fertilizer use. For the period before 1961, we applied the temporal trends from this phosphate rock dataset uniformly across all countries, and adjusted to the country specific 1961 fertilizer estimate. This improved approach for estimating P fertilizer contribution is detailed in the revised manuscript at line 265 under section 2.3.2 of the revised manuscript which reads as follows:

Regarding the time period of 1850 − 1960, when country-level P fertilizer data from FAOSTAT were unavailable, we utilized the temporal dynamics from Cordell et al. (2009) that provides global estimates of phosphate rock production during 1800 − 2000. These estimated P inputs were normalized to align with FAOSTAT data starting in 1961, using 1961 as a reference year for consistency. The global temporal dynamics was then applied across all countries in our study domain for 1850–1960, proportionally scaling the values based on each country's 1961 estimate. This approach allowed *us to generate a temporally coherent dataset, using global phosphate rock production as a proxy for P inputs from fertilizer during the period of limited data availability. The completed annual country-level fertilizer data are referred to as* $Pfer_{soil}(u, y_{1850-2019})$ *(*$kg\ yr^{-1}$*).*

We thank the reviewer for bringing this issue to our attention. We believe the new approach and the adjustments we did have significantly enhances the reliability our P surplus dataset. This revision addresses the concerns you raised, and we are confident it strengthens the methodological integrity of our dataset.

1.4 — I also doubt the assumption of equal distribution rates of treated manure on cropland and pasture. Are there any survey data or studies that can support it?

Reply:

Thank you for your insightful comment. We share the reviewer's critical perspective on this aspect, particularly regarding the broader question of the best approach to estimate distribution rates of treated manure across croplands and pastures. Indeed, this remains an important area and it has been broadly consider as uncertain (Zhang et al., 2021). Different approaches have been reported in literature for estimating treated manure application rates in light of lack of detailed survey data. Accordingly, we applied three different methods – one of them is based on the equal distribution rates of treated manure between cropland and pasture following (Xu et al., 2019) based on generalized assumptions to account for the absence of detailed, consistent data. While such assumptions may not fully reflect actual land management practices, the inclusion of the equal distribution assumption in our study is intended to offer a complementary perspective within our ensemble of methodologies rather than asserting it as the most reliable approach. This assumption has the advantage of dynamically adjusting manure distribution based on changing cropland and pasture areas over time and by country, providing insights into how shifts in land use patterns might influence P dynamics in the long term. We recognize that this approach may carry greater uncertainties than other, more specific country-level methods, and we do not assume that it is preferable to other distribution methodologies.

To address these uncertainties, we further employed an ensemble of approaches that incorporate more granular, country-specific data where available. For instance, the national manure distribution ratios from Ludemann et al. (2023) provide detailed, fixed proportions for each country, while the time-varying national N-based distribution from Einarsson et al. (2021) (used as a proxy for P distribution) incorporates animal type-specific data, refining manure allocation further. These proportions are shown in Figure R8. By integrating these varied methods, our uncertainty-aware approach provides a range of possible outcomes, ensuring that no single assumption, including the equal distribution approach, dominates or unduly influences uncertainty estimates. Instead, our aim is to provide multiple scenarios to capture the inherent uncertainties in manure distribution to cropland and pasture across Europe. We have revised the manuscript at the end of section 2.3.8 to clarify this approach, emphasizing that we do not assert one method as definitively more accurate than another but rather use a suite of methodologies to account for uncertainty comprehensively. The revised text in line 484 reads as follows:

Overall, by integrating two distinct data sources ((FAOSTAT, 2022) and Einarsson et al. (2021)) alongside three manure distribution methods between croplands and pastures, we developed six separate gridded manure estimates for our database. These estimates reflect the uncertainties in our reconstruction, which arise from the selection of different underlying datasets and distribution methods. Each method captures distinct aspects of manure allocation: the equal distribution assumption adjusts dynamically with cropland and pasture area changes over time, while the country-specific ratios from Ludemann et al. (2023) apply fixed national-level allocations. The third method, based on Einarsson et al. (2021), utilizes time-varying nitrogen-based proportions as a proxy for P manure distribution. Supplementary Figure S1 illustrates these proportion of animal manure allocated to cropland and pasture under each method, highlighting the differences and capturing the uncertainties embedded in our approach. By combining these varied assumptions, our *estimates provide a comprehensive view of manure distribution across cropland and pasture, supporting a nuanced analysis of P surplus uncertainty.*

Further, in support of our response, we have included the figure R8 (which corresponds to Supplementary Figure S1) illustrating the proportion of animal manure allocated to cropland and pasture under each of the three methodologies applied in our study. This visual comparison highlights the distinct allocation patterns produced by each method.

1.5 — The calculation of P removal from pasture is very simple. Temporal change of PUE can impact results too. Other impact factors, such as climate, were also not considered.

Reply:

Thank you for your valuable comment regarding the simplicity of our approach to calculating phosphorus (P) removal from pasture. We acknowledge the importance of considering temporal changes in phosphorus use efficiency (PUE) and the potential influence of additional factors like climate on P removal estimates.

Given the inherent complexity of fully estimating P removal from pasture that includes data on changing livestock intake activities, management activities, and other forage activities, in this study we adopted an approach from prior research (Bouwman et al., 2005, 2009; Kaltenegger et al., 2021), utilizing fixed P removal coefficients to approximate PUE for pasture. To capture uncertainty, we used two assumptions: (1) a general P removal coefficient of 0.6, derived from Bouwman et al. (2005), and (2) region-specific PUE values as proxies, based on nitrogen use efficiency (NUE) data from Kaltenegger et al. (2021), with coefficients of 0.4 for Eastern Europe and 0.5 for Western Europe. These approaches allowed us to develop two distinct datasets for P removal from pasture, each reflecting PUE estimates though constant in time. This approach, while simplified, enables an assessment of P removal from pasture in the absence of detailed historical data.

However, we agree that PUE is not static and can vary over time due to changes in agricultural practices, pasture management, and environmental factors, all of which can influence the accuracy of P removal estimates. Likewise, climate factors such as precipitation, temperature, and soil moisture directly affect pasture growth and, consequently, P uptake and cycling. To address the reviewer's comment, we have included a more detailed discussion of these points for further improvements of our datasets in the manuscript. Specifically, we have emphasized that the temporal variability in PUE and climate impacts, such as droughts or temperature extremes, could lead to under- or overestimations of P removal in our dataset. This point under the section 4 at line 870 of the revised manuscript now reads as follows:

Further limitations include the simplification of parameters that likely vary across space and time. While the coefficients used were based on prior research (Bouwman et al., 2005; Kaltenegger et al., 2021), PUE can be highly variable across regions and management practices (Lun et al., 2018; Chowdhury and Zhang, 2021). Our use of fixed coefficients may not fully capture this variability, especially in countries with varying level of grazing intensities or grassland management practices. Furthermore, climate-related factors such as changes in precipitation, temperature, and soil moisture directly affect pasture productivity and thus P uptake and cycling (Martins-Noguerol et al., 2023), which our static approach does not fully encompass. This simplification was necessary due to the lack of detailed historical agricultural records but introduces some degree of uncertainty in our P removal from pasture areas.

Moreover, while dynamic PUE datasets and models, such as those that factor in climate variability (Ijaz et al., 2017) or region-specific grazing practices (Anderson et al., 2020), provide more accurate P removal estimates, they are limited in temporal and spatial scope and do not cover the long historical period of this study. Future work can therefore prioritize integrating such dynamic models where possible. Incorporating these time-varying factors would undoubtedly improve the robustness of our estimates, and we appreciate your suggestion as a pathway for future refinement. Regarding this, we have added the following text in section 5 at line 965 of the revised manuscript:

Another future enhancement would involve refining parameters to account for temporal and spatial variability. For example, crop-specific P uptake rates vary with soil quality, crop variety, and management practices, while pasture P removal is influenced by phosphorus use efficiency (PUE) and meteorological and hydrological variables such as precipitation, temperature, and soil moisture. Future work should incorporate dynamic PUE estimates to better capture time-varying removal rates driven by regional grazing practices, crop types, and changing weather patterns.

1.6 — Since there are so many weaknesses in the method, I strongly suggest adding one section of the limitation of this data.

Reply: Thank you for your feedback. In response to your suggestion, we have added two dedicated sections — *4. Potential use and limitations of the dataset* and *5. Directions for future improvement of the dataset* — in the revised manuscript. These sections thoroughly outline the key methodological constraints of our dataset. By explicitly addressing the limitations and identifying directions for future improvement, we aim to enhance transparency about potential uncertainties in our findings and provide a balanced perspective on the dataset's strengths and limitations. We believe these additions will offer readers valuable insights into both the current applicability of the dataset and directions for its continued refinement.

References

Anderson, K. R., Moore Jr, P. A., Pilon, C., Martin, J. W., Pote, D. H., Owens, P. R., Ashworth, A. J., Miller, D. M., and DeLaune, P. B.: Long-term effects of grazing management and buffer strips on phosphorus runoff from pastures fertilized with poultry litter, Tech. rep., Wiley Online Library, 2020.

Antikainen, R., Haapanen, R., Lemola, R., Nousiainen, J. I., and Rekolainen, S.: Nitrogen and phosphorus flows in the Finnish agricultural and forest sectors, 1910–2000, Water, air, and soil pollution, 194, 163–177, 2008.

Bartholomé, E. and Belward, A. S.: GLC2000: A new approach to global land cover mapping from earth observation data, Int. J. Remote Sens., 26, 1959–1977, https://doi.org/10.1080/01431160412331291297, 2005.

Bouwman, A. F., Van Brecht, G., and Van Der Hoek, K. W.: Global and regional surface nitrogen balances in intensive agricultural production systems for the period 1970-2030, Pedosphere, 15, 137–155, 2005.

Bouwman, A. F., Beusen, A. H., and Billen, G.: Human alteration of the global nitrogen and phosphorus soil balances for the period 1970-2050, Global Biogeochemical Cycles, 23, https://doi.org/10.1029/2009GB003576, 2009.

Chowdhury, R. B. and Zhang, X.: Phosphorus use efficiency in agricultural systems: A comprehensive assessment through the review of national scale substance flow analyses, Ecological Indicators, 121, 107 172, 2021.

Cordell, D., Drangert, J.-O., and White, S.: The story of phosphorus: global food security and food for thought, Global environmental change, 19, 292–305, 2009.

Di Gregorio, A.: Land cover classification system: classification concepts and user manual: LCCS, vol. 2, Food & Agriculture Org., 2005.

Einarsson, R., Sanz-co, A., Aguilera, E., Billen, G., Josette, G., G, H. J. M. V., and Lassaletta, L.: Crop production and nitrogen use in European cropland and grassland 1961 – 2019, Scientific Data, pp. 1–29, https://doi.org/10.1038/s41597-021-01061-z, 2021.

FAOSTAT: FAOSTAT (Food and Agriculture Organization Corporate Statistical Database): Livestock Manure, available at:, https://www.fao.org/faostat/en/#data/EMN, [Accessed: 01-May-2022], 2022.

Guejjoud, H., Curie, F., and Grosbois, C.: Analyzing a century of agricultural phosphorus surplus and its long-term key drivers in France, Nutrient Cycling in Agroecosystems, https://doi.org/10.1007/s10705-023-10300-1, 2023.

Holland, E. A., Lee-Taylor, J., Nevison, C., and Sulzman, J.: Global N Cycle : Fluxes and N 2 O Mixing Ratios Originating from Human Activity, Report, p. 4, 2005.

Ijaz, W., Ahmed, M., Asim, M., and Aslam, M.: Models to study phosphorous dynamics under changing climate, Quantification of climate variability, adaptation and mitigation for agricultural sustainability, pp. 371–386, 2017.

Kaltenegger, K., Erb, K. H., Matej, S., and Winiwarter, W.: Gridded soil surface nitrogen surplus on grazing and agricultural land: Impact of land use maps, Environ. Res. Commun., 3, https://doi.org/10.1088/2515-7620/abedd8, 2021.

Ludemann, C. I., Wanner, N., Chivenge, P., and ...: A global reference database in FAOSTAT of cropland nutrient budgets and nutrient use efficiency: nitrogen, phosphorus and potassium, 1961–2020, Earth System Science Data, pp. 1–24, https://essd.copernicus.org/preprints/essd-2023-206/%0Ahttps://essd.copernicus.org/preprints/essd-2023-206/essd-2023-206.pdf, 2023.

Lun, F., Liu, J., Ciais, P., Nesme, T., Chang, J., Wang, R., Goll, D., Sardans, J., Peñuelas, J., and Obersteiner, M.: Global and regional phosphorus budgets in agricultural systems and their implications for phosphorus-use efficiency, Earth System Science Data, 10, 1–18, 2018.

Martins-Noguerol, R., Moreno-Pérez, A. J., Pedroche, J., Gallego-Tévar, B., Cambrollé, J., Matías, L., Fernández-Rebollo, P., Martínez-Force, E., and Pérez-Ramos, I. M.: Climate change alters pasture productivity and quality: Impact on fatty acids and amino acids in Mediterranean silvopastoral ecosystems, Agriculture, Ecosystems & Environment, 358, 108 703, 2023.

Muntwyler, A., Panagos, P., Pfister, S., and Lugato, E.: Assessing the phosphorus cycle in European agricultural soils: Looking beyond current national phosphorus budgets, Science of the Total Environment, 906, 167 143, 2024.

Oenema, O., Kros, H., and de Vries, W.: Approaches and uncertainties in nutrient budgets: implications for nutrient management and environmental policies, European Journal of Agronomy, 20, 3–16, 2003.

Panagos, P., Muntwyler, A., Liakos, L., Borrelli, P., Biavetti, I., Bogonos, M., and Lugato, E.: Phosphorus plant removal from European agricultural land, Journal fur Verbraucherschutz und Lebensmittelsicherheit, 17, 5–20, https://doi.org/10.1007/s00003-022-01363-3, 2022.

Ringeval, B., Demay, J., Goll, D. S., He, X., Wang, Y. P., Hou, E., Matej, S., Erb, K. H., Wang, R., Augusto, L., Lun, F., Nesme, T., Borrelli, P., Helfenstein, J., McDowell, R. W., Pletnyakov, P., and Pellerin, S.: A global dataset on phosphorus in agricultural soils, Scientific Data, 11, 1–34, https://doi.org/10.1038/s41597-023-02751-6, 2024.

Sarrazin, F. J., Attinger, S., and Kumar, R.: Gridded dataset of nitrogen and phosphorus point sources from wastewater in Germany (1950–2019), Earth System Science Data, 16, 4673–4708, 2024.

Van Meter, K. J., McLeod, M. M., Liu, J., Tenkouano, G. T., Hall, R. I., Van Cappellen, P., and Basu, N. B.: Beyond the Mass Balance: Watershed Phosphorus Legacies and the Evolution of the Current Water Quality Policy Challenge, Water Resources Research, 57, e2020WR029 316, https://doi.org/https://doi.org/10.1029/2020WR029316, 2021.

Xu, R., Tian, H., Pan, S., Dangal, S. R., Chen, J., Chang, J., Lu, Y., Skiba, U. M., Tubiello, F. N., and Zhang, B.: Increased nitrogen enrichment and shifted patterns in the world's grassland: 1860–2016, Earth System Science Data, 11, 175–187, 2019.

Zhang, X., Zou, T., Lassaletta, L., Mueller, N. D., Tubiello, F. N., Lisk, M. D., Lu, C., Conant, R. T., Dorich, C. D., Gerber, J., Tian, H., Bruulsema, T., Maaz, T. M. C., Nishina, K., Bodirsky, B. L., Popp, A., Bouwman, L., Beusen, A., Chang, J., Havlík, P., Leclère, D., Canadell, J. G., Jackson, R. B., Heffer, P., Wanner, N., Zhang, W., and Davidson, E. A.: Quantification of global and national nitrogen budgets for crop production, Nat. Food, 2, 529–540, https://doi.org/10.1038/s43016-021-00318-5, 2021.

Zou, T., Zhang, X., and Davidson, E.: Global trends of cropland phosphorus use and sustainability challenges, Nature, 611, 81–87, 2022.

List of Figures

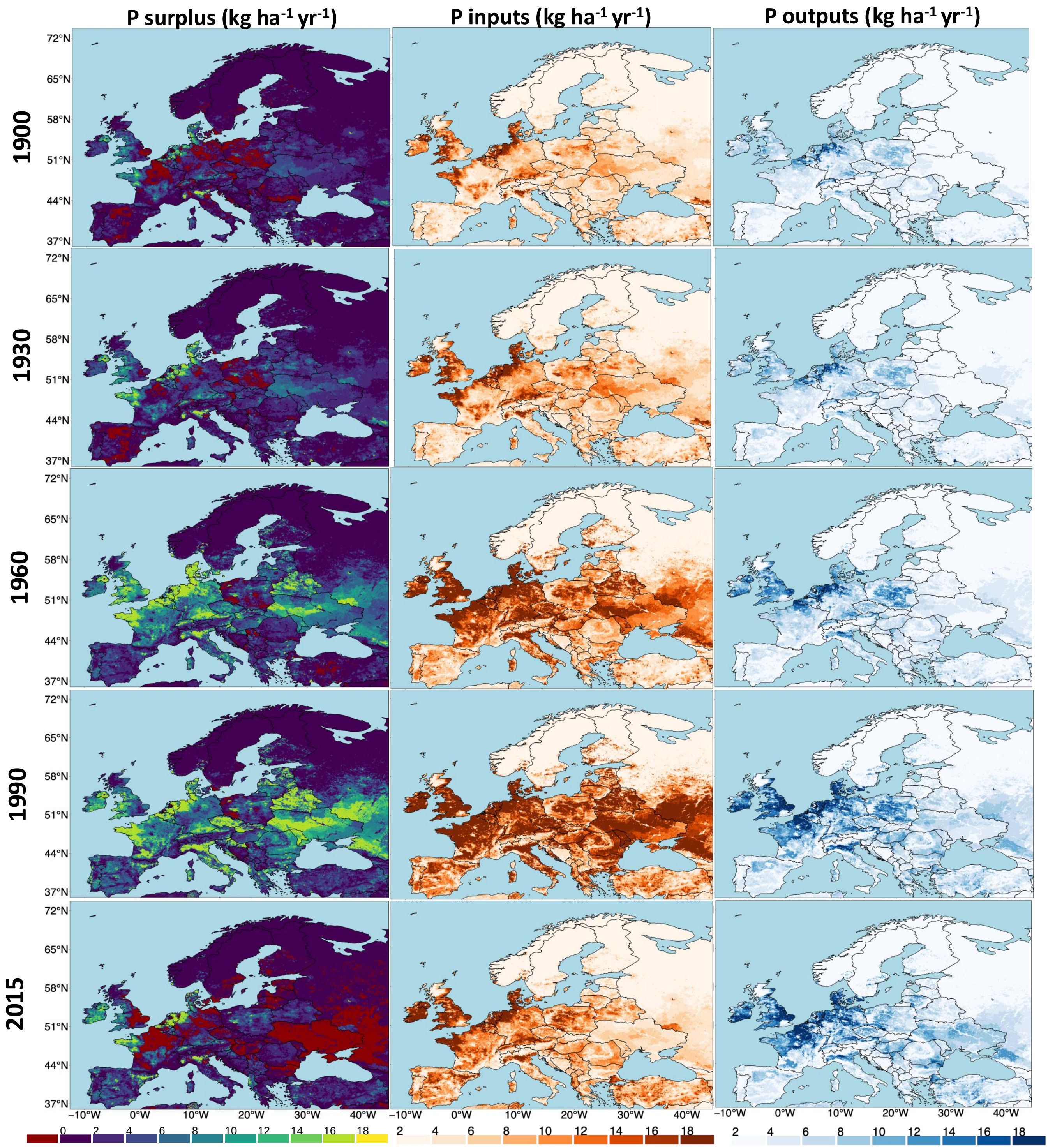

Figure R1. Snapshots of P surplus, P inputs and P outputs ($kg\ ha^{-1}$ of grid physical area yr^{-1}) across Europe. The figure shows the annual spatial variation in P surplus, P inputs and P outputs given as the mean of our 48 P surplus, P inputs and P outputs estimates for the selected years.

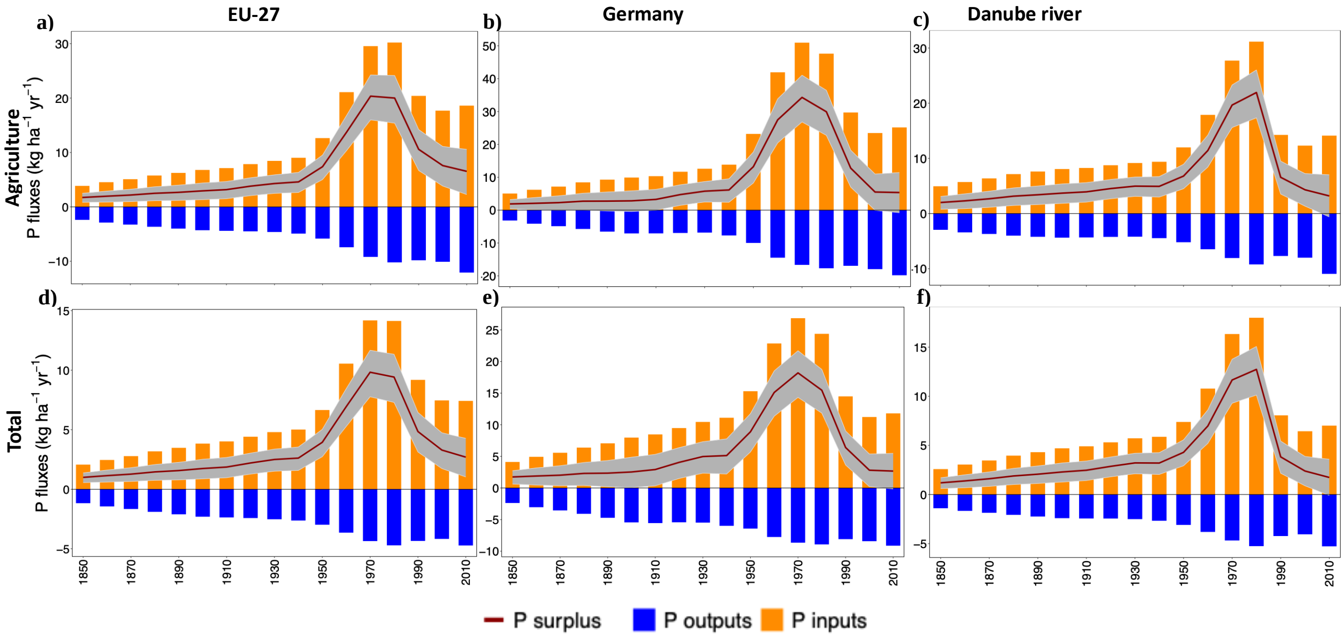

Figure R2. Decadal trajectory of agricultural P surplus ($kg\ ha^{-1}$ of agricultural area yr^{-1}) and total P surplus ($kg\ ha^{-1}$ of physical area yr^{-1}) and its contributing components for the EU-27, Germany, and the Danube river basin from 1850 to 2019. Upward orange bars represent the average of 48 P inputs, while downward blue bars indicate the average of 48 P outputs, showing decadal means. The grey ribbon shows the range (min and max) of the 48 P surplus estimates, with the red line representing the average value for each decade. **(a-c)** Agricultural P surplus for EU-27, Germany and Danube river, **(d-f)** Total P surplus for EU-27, Germany and Danube river

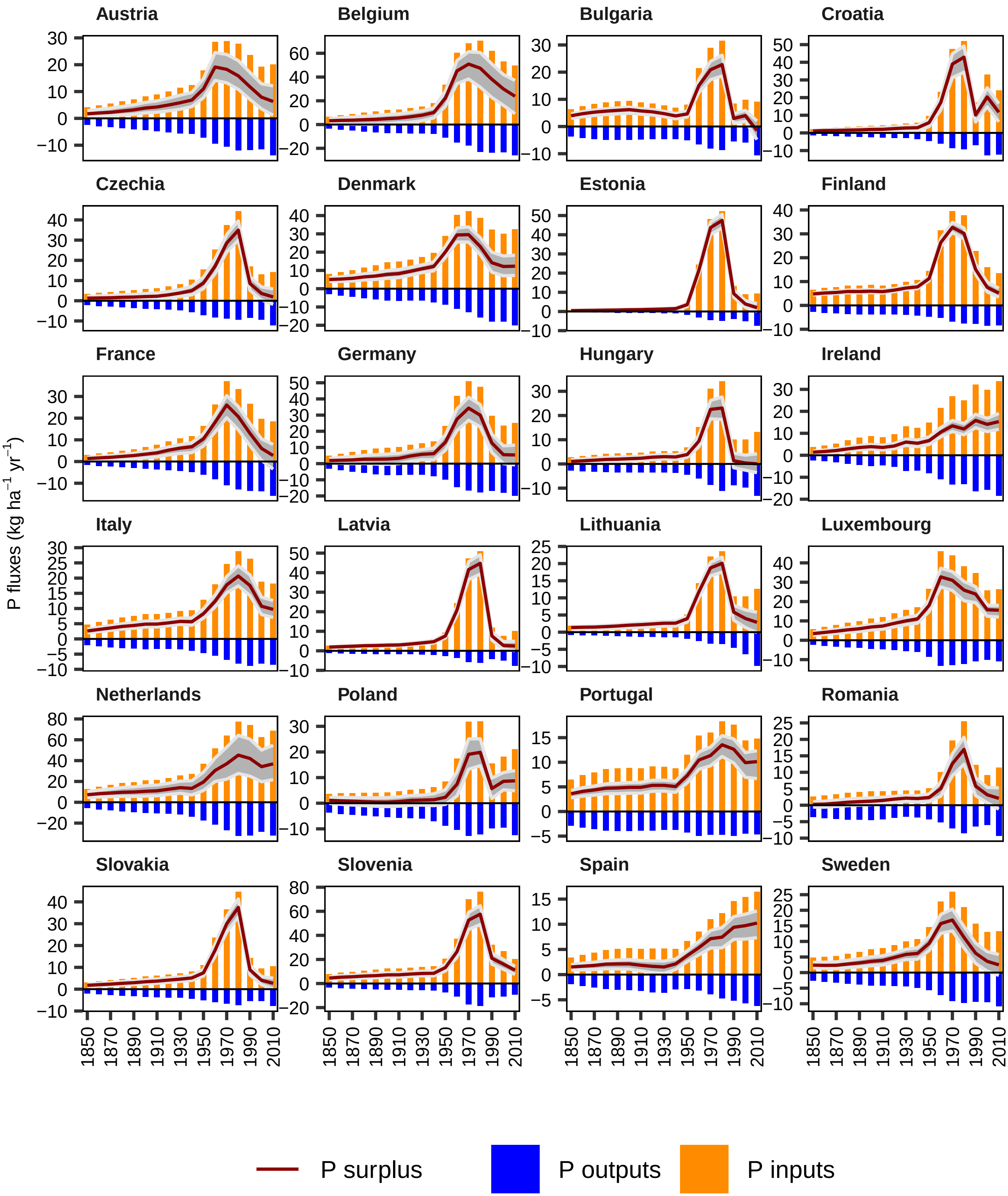

Figure R3. Decadal trajectory of agricultural P surplus ($kg\ ha^{-1}$ of agricultural area yr^{-1}) and its contributing components for different European countries from 1850 to 2019. Upward orange bars represent the average of 48 P inputs, while downward blue bars indicate the average of 48 P outputs, showing decadal means. The grey ribbon shows the range (min and max) of the 48 P surplus estimates, with the red line representing the average value for each decade.

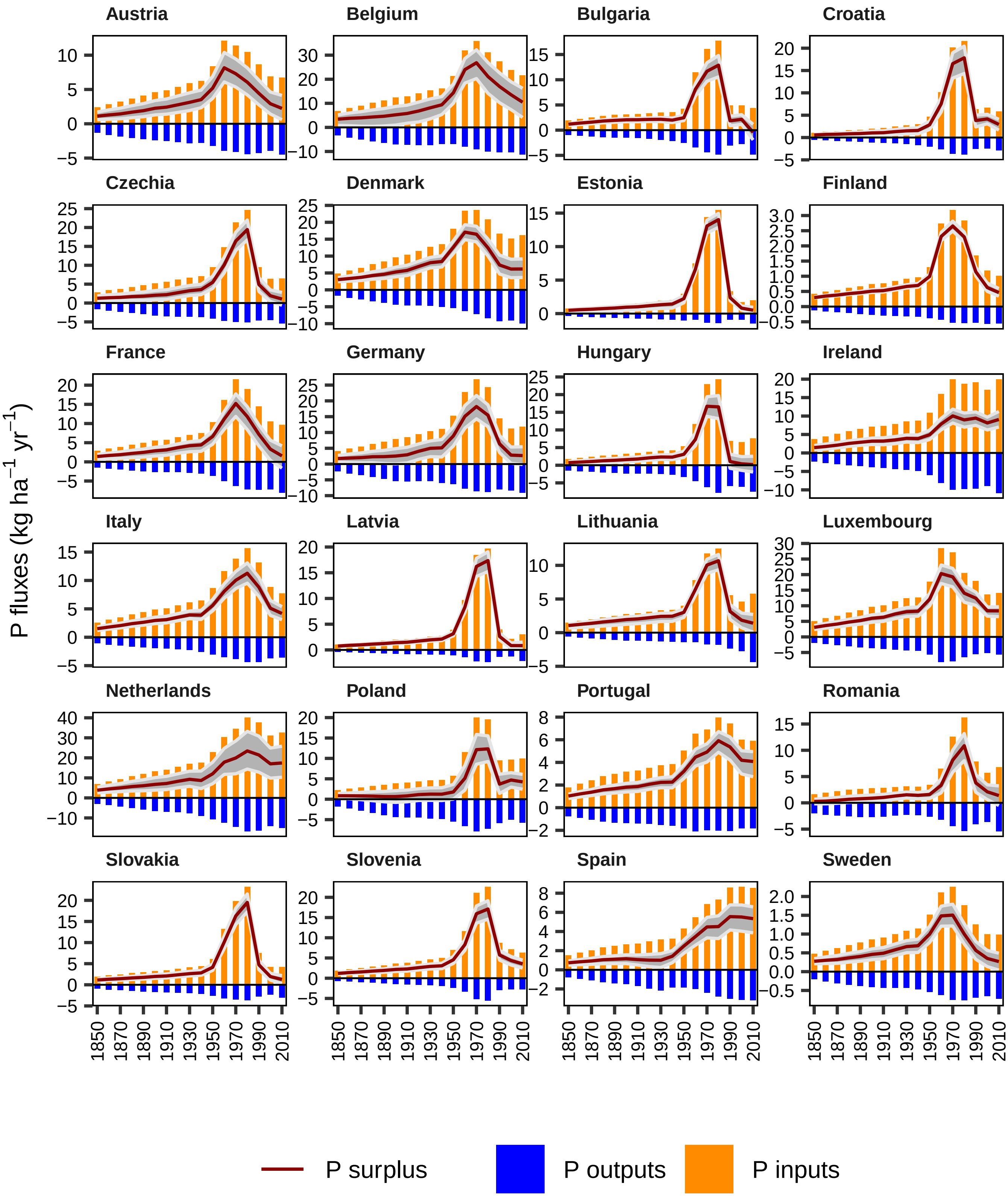

Figure R4. Decadal trajectory of total P surplus ($kg\ ha^{-1}$ of physical area yr^{-1}) and its contributing components for different European countries from 1850 to 2019. Upward orange bars represent the average of 48 P inputs, while downward blue bars indicate the average of 48 P outputs, showing decadal means. The grey ribbon shows the range (min and max) of the 48 P surplus estimates, with the red line representing the average value for each decade.

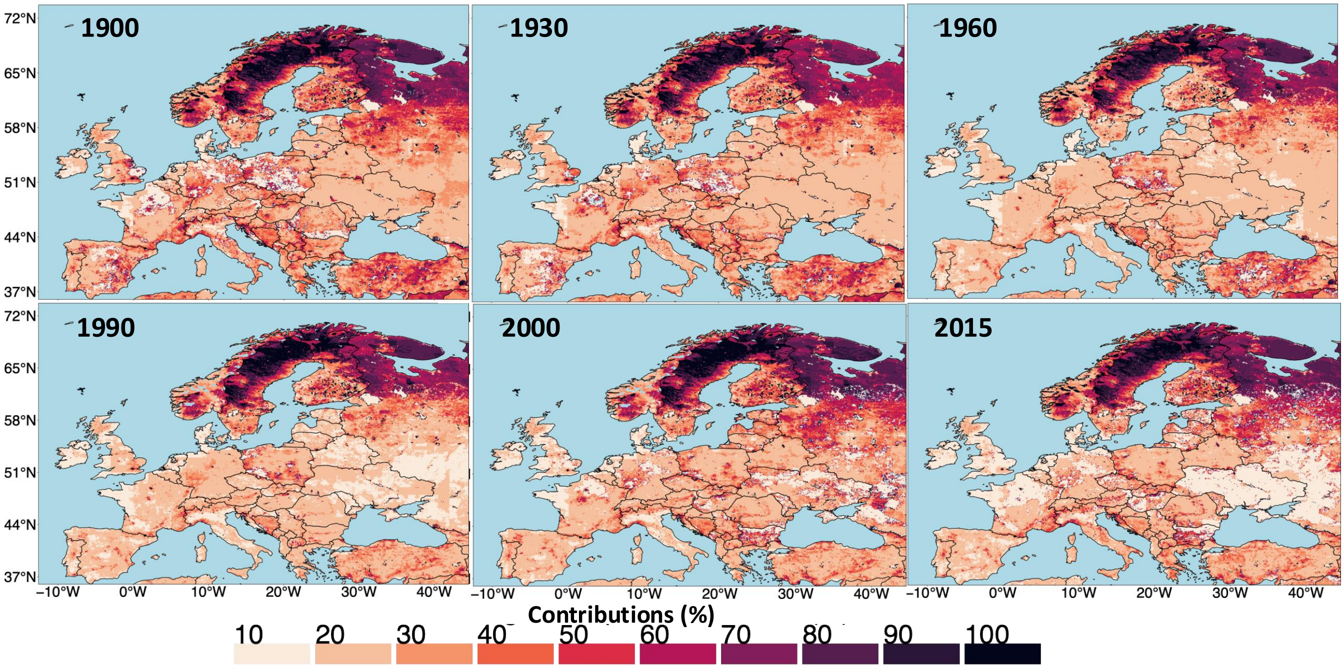

Figure R5. Snapshots showing the spatial distribution of the contribution (%) of non-agricultural P surplus to the total P surplus across Europe for selected years. The figure highlights the annual variation in the proportion of non-agricultural P surplus to the total P surplus (averaged from 48 P surplus estimates) across different regions, illustrating the evolving role of non-agricultural sources in European P dynamics over time.

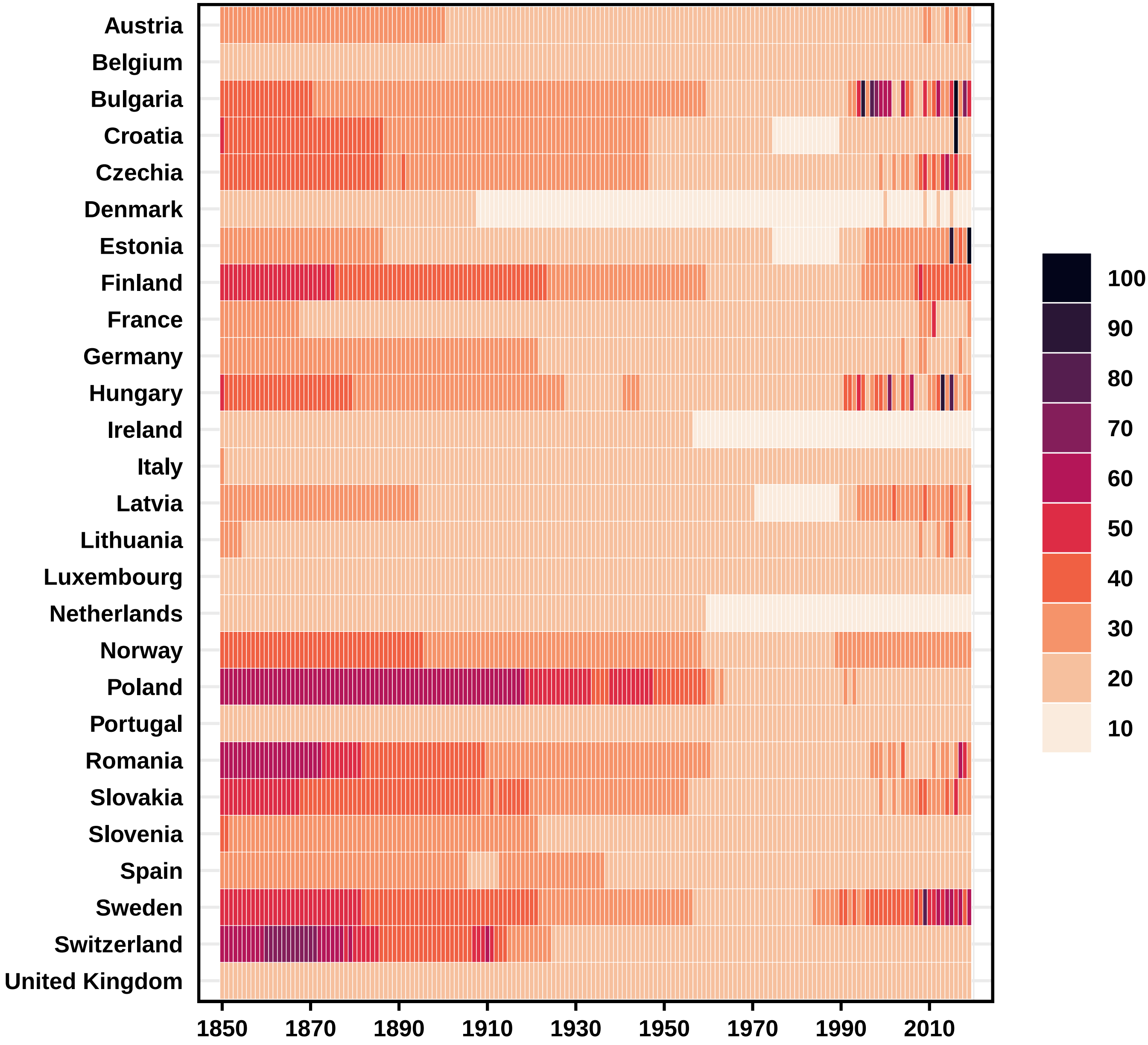

Figure R6. Heat map showing the temporal evolution of the contribution (%) of non-agricultural P surplus to the total P surplus across different European countries from 1850-2019. The figure highlights the annual variation in the proportion of non-agricultural P surplus to the total P surplus (averaged from 48 P surplus estimates) across different European countries, illustrating the evolving role of non-agricultural sources in European P dynamics over time. In recent years, countries such as Hungary, Bulgaria, Estonia and Croatia have recorded peak values, which is mainly due to a lower agricultural P surplus, as a result of which the relative share of the non-agricultural P surplus in the total P surplus has amplified.

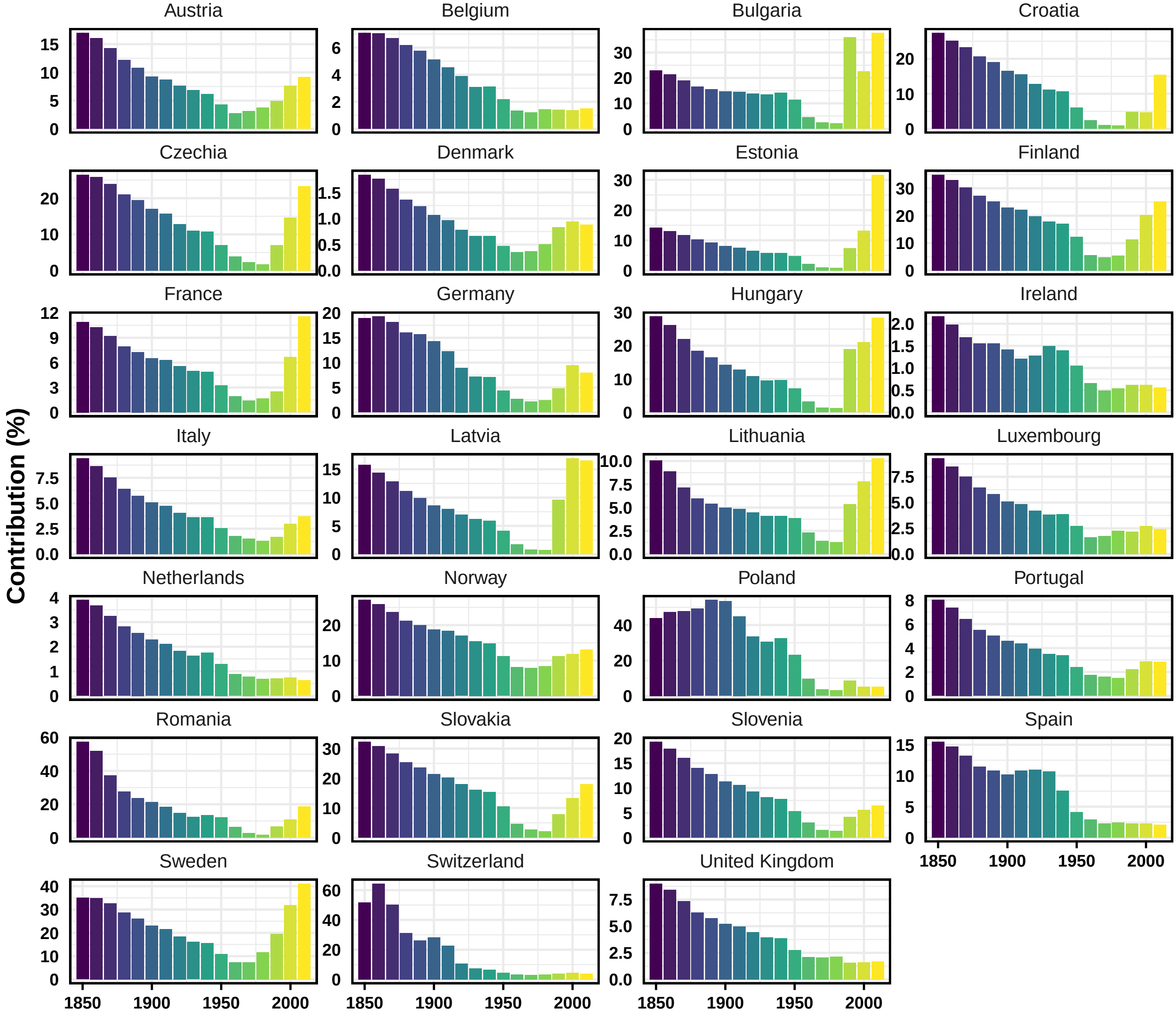

Figure R7. Decadal contribution of non-agricultural P surplus to the total P surplus across different European countries

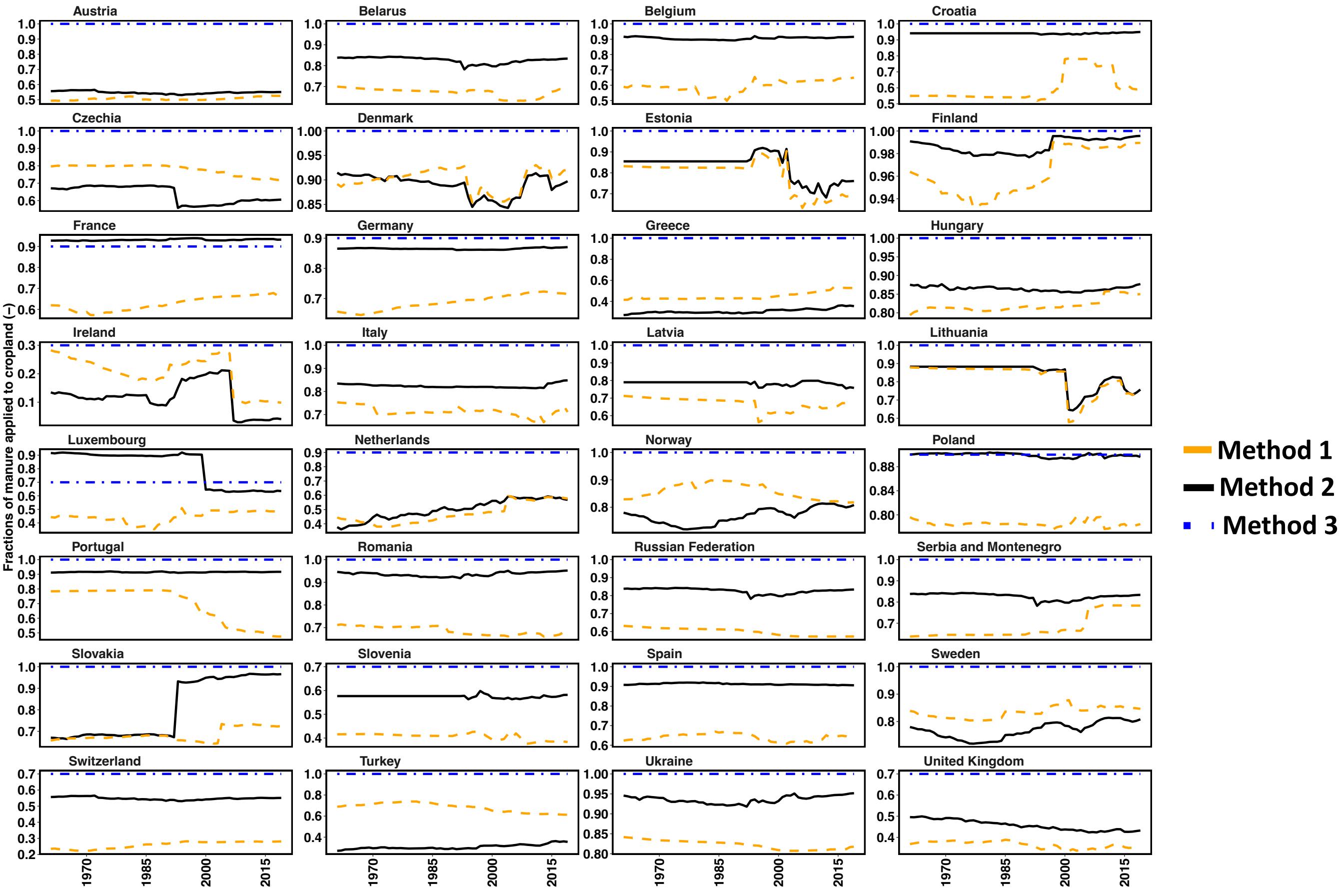

Figure R8. Fractions of manure distribution to cropland based on different methods utilized in this study. Method 1 represents the fraction of distribution of animal manure to cropland based on the equal distribution rates for cropland and pasture within each grid cell. Method 2 shows the fraction using the time-varying national proportions of nitrogen (N) manure applied to both cropland and pasture, as provided by Einarsson et al. (2021). Method 3 shows the manure distribution based on country-level data on manure application proportions to cropland and pasture, as reported by Ludemann et al. (2023).