Metrics and methods for characterizing dairy farm intensification using farm survey data

Gonzalez Mejia, Alejandra; Styles, David; Wilson, Paul; Gibbons, James

PLoS ONE

DOI:
10.1371/journal.pone.0195286

Published: 09/05/2018

Peer reviewed version

PLOS ONE

Metrics and methods for characterizing dairy farm intensification using farm survey data

--Manuscript Draft--

Manuscript Number:	PONE-D-17-37318R2
Article Type:	Research Article
Full Title:	Metrics and methods for characterizing dairy farm intensification using farm survey data
Short Title:	Metrics and methods for characterizing dairy farm intensification using farm survey data
Corresponding Author:	James Gibbons Bangor University Bangor, UNITED KINGDOM
Keywords:	Model-based clustering; PCA; farm typology; intensification; dairy systems; key performance indicators; Trend analysis
Abstract:	Evaluation of agricultural intensification requires comprehensive analysis of trends in farm performance across physical and socio-economic aspects, which may diverge across farm types. Typical reporting of economic indicators at sectorial or the "average farm" level does not represent farm diversity and provides limited insight into the sustainability of specific intensification pathways. Using farm business data from a total of 7281 farm survey observations of English and Welsh dairy farms over a 14-year period we calculate a time series of 16 key performance indicators (KPIs) pertinent to farm structure, environmental and socio-economic aspects of sustainability. We then apply principle component analysis and model-based clustering analysis to identify statistically the number of distinct dairy farm typologies for each year of study, and link these clusters through time using multidimensional scaling. Between 2001 and 2014, dairy farms have largely consolidated and specialized into two distinct clusters: more extensive farms relying predominantly on grass, with lower milk yields but higher labour intensity, and more intensive farms producing more milk per cow with more concentrate and more maize, but lower labour intensity. There is some indication that these clusters are converging as the extensive cluster is intensifying slightly faster than the intensive cluster, in terms of milk yield per cow and use of concentrate feed. In 2014, annual milk yields were 6,835 and 7,500 l/cow for extensive and intensive farm types, respectively, whilst annual concentrate feed use was 1.3 and 1.5 tonnes per cow. For several KPIs such as milk yield the mean trend across all farms differed substantially from the extensive and intensive typologies mean. The indicators and analysis methodology developed allows identification of distinct farm types and industry trends using readily available survey data. The identified groups allow the accurate evaluation of the consequences of the reduction in dairy farm numbers and intensification at national and international scales.
Order of Authors:	Alejandra Gonzalez Mejia
	David Styles
	Paul Wilson
	James Gibbons
Opposed Reviewers:	
Response to Reviewers:	As in the previous submission we thank both reviewers for their helpful efforts and input. For clarity in the point-by-point response we have deleted any comments and responses that relate to the previous revision and retained only those where issues remained. Reviewer #1's suggestions were all suggested edits and we have made all the changes suggested and reread through the manuscript for English as suggested

Reviewer #2's suggestions while recognising that the manuscript has been improved asks for further statistical explanation as well as suggested edits. We have done all the suggested edits and further explained the statistical approach. The main addition is Figure 1 which shows the workflow of the statistical analysis. This was a great suggestion and really improves the manuscript. We have also added further written explanation to the Data Analysis section. Reviewer #2 also asked that we illustrate the methods with a simplified data set, for space and emphasis reasons we have not done this within the manuscript but point to a book with a freely available PDF version: https://web.stanford.edu/~hastie/ElemStatLearn/
that illustrates the approaches with simple data sets.

Reviewer #1
General comments
The manuscript has improved but it needs some other changes. In my opinion, it can be published after minor revision.
Firstly I consider very important to improve the aim and the conclusion according to the new title
Moreover, after the first revision, some new or revised sentences rather unclear and several typos appeared. I suggest rereading accurately the entire manuscript to amend the errors and clarify the meaning of some phrases.
Finally some comments were only partially addressed

Specific comments

Title
Thank you for having modified the title. I can suggest further optional alternatives:
1) Metrics for characterizing Dairy Farm Intensification: the case of England & Wales Between 2001 And 2014
2) Metrics for characterizing Dairy Farm Intensification
This is because the methods proposed is also applicable to other time periods

Thank you for the suggestion you make a good point about time so we have changed to "Metrics and methods for characterizing dairy farm intensification using farm survey data"

Abstract
Line 10: Please, replace the 2nd trend with a synonym in the first sentence: Evaluation of the sustainability of trends in agricultural intensification requires comprehensive analysis of trends in farm performance across physical and socio-economic aspects

Changed the first to "changes"

Introduction
Line 35: 5% of global anthropogenic gas emissions?
Done

Lines 38-39: The sentence is not so clear.
Changed to "One route to this is to reduce land-use intensity of milk production by increasing milk yields per cow (2)."

Lines 39-41: I do not understand the meaning of the sentence

Changed to "However, without advances in technology an environmental gain will only be achieved if the increase in production per cow out paces the increase in demand."

Lines 44-46: I suggest reversing the sentences: The UK dairy industry is 10th largest global producer of cow milk (accounting for 2.2% of world production) (5) and is an exemplar of worldwide intensification trends.
Done

Line 59: I suggest avoiding double round brackets. You can use a comma or a semicolon before the number of the citation

I would prefer to keep as it consistently differentiates a literal number from a citation, however we are happy to have this formatted as per editor or journal style preferences if this is required.

Line 61: because despite. Are they both useful?
Inserted a comma between “because, despite”

Lines 117-127: please rewrite and refine the aim of the study
This is methods rather than the aim but we have removed “ratios, densities and intensities”

Methods
Lines 91-93: I suggest conjugating verbs to the past simple (identified…investigated)
Done

Line 125: to provide
Done

Results
I still see a lack of consistency among the names of indicators in table 1, fig. 1, fig 3, In the text and now also in fig s1 in the supplementary material.
Thanks, made much more consistent, though still retained some extra words in Table 1 to avoid ambiguity when the units are not presented

Line 185: I suggest deleting the reference to fig 1 because it has just been recalled in the previous line
Done

Fig. 1
In the figure title please put “Key performance indicators” in extenso instead of KPI
Done

Despite the improvements in the readability of indicators names in the PCAs, I can still count only 15 indicators in the first two PCAs and 16 in the third. The indicator Non-cash crop area ha ha-1 seems to appear only in the last PCA
Thanks for the reminder about this, the scale on the axis was omitting the final variable, now adjusted so the scale is -0.45 to 0.45 on all three graphs.

Table 3
Replacement rate seems in bold and/or in a different font.
Well spotted, 10 point rather than 9 point, fixed

Fig 3
Please, check the English of the figure title
In the figure title I suggest putting “Key performance indicators (KPI)” in extenso instead of KPI, the first time
Changed to “Fig 4 – Trends in mean Key Performance Indicator values for all identified clusters over the period 2001 – 2014. The number of farms in each cluster is represented by the size of symbol. Intensive systems are represented by triangles and extensive systems by circles. The solid black line represents the KPI annual average. The distance among all clusters in all years of study is represented by the colour scale MDS. This distance allows identifying which clusters are more similar.”

Discussion
Lines 293-296: Check the English of the phrase and, if possible, break into shorter sentences.
Changed end to “if these businesses choose not to be surveyed.“

Lines 345: avoid if possible the repetition of capture
Changed the second to “include”

Lines 349-352: Check the English of the phrase
Removed “that”

Conclusions
I think conclusion section needs an improvement. The conclusion focuses only on the specific results of the England and Wales case. Conclusion about the method and its contribution to research are missing.
Added “In the method developed here indicators from farm business survey data coupled with robust clustering identify groups of farms and trends over time.”

Line 381: I suggest adding the specification “in England and Wales”
Done

Reviewer #2: I would like to thank the authors for their efforts. This new version of the manuscript has been significantly improved. However, I am afraid that the manuscript is still not ready for publication, because the methodology remains insufficiently explained. Answers to many of my comments are incomplete or have been completely ignored. I have made great efforts in this and the previous version to help the authors improve the presentation of their methods and would like to see my recommendations materialize. I think that this study is very interesting so I am only asking the authors to take my recommendations into account to make the manuscript more accessible to readers.

Please find my comments below. Every comment starts with ‘RESPONSE:’

l.66: Sentence starting with ‘While’ does not make sense. I think you meant to link this sentence with the previous one as follows: ‘[…] (8), while Alvarez et al. (9) […]’. This would result to a very long sentence. I recommend that you just start the second sentence as ‘Alvarez et al. (9) […]’.
Done

l:59:82: RESPONSE: Thank you for the edits. Choice of French beef example fair enough. Only comment here is that I would like to see a better linking between the following sentences: ‘It might be expected that intensification of dairy farms will result in more efficient farms growing at the expense of less efficient farms. However, evidence from the livestock sector suggests that farms may become less economically efficient because, despite an increase in investment in capital, technology and concentrate feed, output may remain constant over several decades, as is the case for French beef farms (Veysset et al., 2015)’.
Inserted “Given this range in efficiency it…”at the start of the 2nd sentence

l.122-123: A comma following ‘e.g.’ or ‘i.e.’ is unnecessary. This typo appears in several places in the manuscript. Please replace ‘e.g.,’ and ‘i.e.,’ with ‘e.g.’ and ‘i.e.’ respectively.
Thanks, eliminated all rogue commas

Statistics. RESPONSE: Thank you for your efforts. However, as with the earlier version, I think that you have not placed enough effort on satisfactorily shaping this subsection and on addressing my comments. For example, you are assuming that everyone is familiar with PCA loadings (l.215) and have not explained what they are. In fact, you have not even explained what PCA is, how it works and why you have used it. You obviously want to represent KPIs with uncorrelated variables. Explain why.
Please briefly describe PCA, it is very easy to do so. See section 4 in Jollands et al. (2004) - an excellent brief description of PCA. Please add textbook references for PCA and Procrustes rotation as no one is going to learn/understand these methods just by reading the R documentation.
Also, it is unclear why Procrustes rotation of the loadings is necessary. That ‘the sign of component loadings is arbitrary’ (l.219-220) does not help the reader understand why Procrustes is used, especially given that many readers may not know what loadings are, and that they have a sign.
With k-means, clusters are not necessarily of ‘equal size’ (l.231)? See Alvarez et al. (2008).
My requests for a diagram illustrating how the different models/methods are combined;

and for a short example demonstrating your novel approach and its advantages ‘in action’ seem to have been completely ignored. Right now, this section feels ‘overloaded’ in terms of methodology: PCA, Procrustes rotation, Gaussian mixture model-based clustering, Expectation Maximization algorithm, Maximum Likelihood, BIC, multidimensional scaling, shape, volume and orientation of multidimensional datasets… this is too much for the ‘intelligent lay reader’. Personally, I am not really following. A visual summary of how all methods combined step-by-step, as well as a trivial example (perhaps visual too), are absolutely necessary.

Sorry that we did not provide more explanation in the last revision. Thank you for the suggestions here, we now include a new Figure (Figure 1) which illustrates the flow through the methods and how they relate. We have further edited the introduction to the Data Analysis section to describe PCA. Thank you for the reference suggestion, we have now inserted a reference to Hastie et al. which is a very good book, a PDF is freely available from the author (https://web.stanford.edu/~hastie/ElemStatLearn/) and it covers the majority the methods; we use the additional reference to the Mardia et al. book that covers the rest.
While illustrating the methods with simple examples is a good suggestion we think that this would change the emphasis of the paper and there would not be space to do this justice. As the Hastie et al. book illustrates the methods with simple and consistent examples we don’t think it is necessary to duplicate this within the manuscript.
We have expanded the explanation of Procrustes rotation to explain further this aspect.
Technically you are correct to state that k-mean and related methods don’t explicitly assume clusters are the same size. However, in the case where they are different sized very surprising cluster results are produced. I.e. practically the assumption holds. There are a couple of useful blog entries on this with illustrations
https://blog.learningtree.com/assumptions-ruin-k-means-clusters/
http://varianceexplained.org/r/kmeans-free-lunch/
Taking on board the recommendations, the main text edits are at the start of the Data Analysis section which now reads:
“We use a suite of statistical methods and workflow to analyse the data as shown in Fig 1. Further details of all the analysis methods with illustrations on simple data sets are available in (43,44) in particular we recommend chapter 14 of Hastie et al. All code to reproduce the data analysis is available on request from the authors. PCA (principal components analysis) was used to explore the relationship among KPIs (i.e. identification of fundamental farm properties) and how these relationships change over time. The usual aim is to reduce multiple dimensions down to two or three for illustration and analysis purposes. PCA creates new linear combinations of existing variables (components) ranked to explain as much variation as possible. The relative weighting of each KPI on each component is then termed the loading and value each farm on the component the score. For the set of KPIs to be a useful measure of farms over time, the relationship between KPIs should be relatively constant but change should result in farms changing their position along the KPI dimensions. PCA was calculated in R (45) and Procrustes rotation of the first 3 KPI loadings identified by PCA was used to compare the structure of each year and compare structure between years with the vegan package (46). The Procrustes analysis rotated the PCA loadings to minimize the sum of squares of the difference in distance between loading for each year pair, a small total sum of squares indicating the relationship between the individual KPIs between years was similar, a large difference that the relationship changed between years. The rotation is necessary to fairly compare between years as the relationship between the variables and hence the relative loadings may remain constant over time but the absolute loadings may change and the sign of component loadings is arbitrary (can be positive or negative depending on the algorithm or data used).”

l.189: ‘[…] and cash crops provide information […]’. Do you want to say ‘[…] and cash crops that provide information […]’?
Inserted “to” to correct the grammar

l.141: Can you please confirm that the assumption of a Gaussian distribution does not impact on the validity of your analysis. How do your KPIs look like? I would recommend that you provide histograms and boxplots of your 16 indicators as supplementary information.

	The shape of the clusters is quite flexible in the sense they can range from speherical to long and thin, though (as you would expect in a cluster) could not be discontinuous or not smooth.. As indicated in Table 2 the majority of clusters were of the 'VVV' type indicating that this flexibility was necessary. RESPONSE: Apologies for my limited understanding, but I do not see how this answers my question about the assumption of normality. Possibly we misunderstood you here in relation to the initial suggestion. The boxplots in the supplementary material included in the last revision did illustrate the distribution of the variables. Non-Gaussian distributions would be accommodated because e.g. in the case of a multimodal distribution, there could be a cluster at each mode in that dimension. I.e individual clusters are multivariate Gaussian though not constrained to be spherical, but this does not constrain the underlying variables to be Gaussian. l.148-154: Can you explain further, and graphically, what is meant by 'shape, volume, orientation'. EVI, VEV and VVV mean little to me as they currently stand. We now clarify in the main text: Line 173-175 "The first identifier denotes volume (equal or variable size), the second shape (spherical or not) and the third orientation (aligned or not)." RESPONSE: See my earlier comments about graphically explaining all this (and all the rest). See above response l.158: At this point, I have a very faint idea how your novel approach does what it does. Please help the reader. We hope that the edits above help with this RESPONSE: Not much. Please add: (i) a diagram illustrating how the different models/methods are combined; and (ii) a short example demonstrating your novel approach and its advantages 'in action'. See earlier comments. As above we have added the diagram and further explanation l.309-322: RESPONSE: Can you please say in the manuscript that you are happy to share your code upon request? Inserted "Note that all code to reproduce the data analysis is available on request from the authors." Towards the start of the Data Analysis section l.464: 'the' is missing. Fixed
Additional Information:	
Question	**Response**
Financial Disclosure Please describe all sources of funding that have supported your work. **This information is required for submission and will be published with your article, should it be accepted.** A complete funding statement should do the following: Include **grant numbers and the URLs** of any funder's website. Use the full name, not acronyms, of funding institutions, and use initials to identify authors who received the funding. **Describe the role** of any sponsors or funders in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.	The study was funded by the Sêr Cymru National Research Network for Low Carbon Energy and Environment (NRN-LCEE). http://www.nrn-lcee.ac.uk/. he funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

If the funders had **no role** in any of the above, include this sentence at the end of your statement: "*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*" However, if the study was **unfunded**, please provide a statement that clearly indicates this, for example: "*The author(s) received no specific funding for this work.*" * typeset	
Competing Interests You are responsible for recognizing and disclosing on behalf of all authors any competing interest that could be perceived to bias their work, acknowledging all financial support and any other relevant financial or non-financial competing interests. Do any authors of this manuscript have competing interests (as described in the PLOS Policy on Declaration and Evaluation of Competing Interests)? **If yes**, please provide details about any and all competing interests in the box below. Your response should begin with this statement: *I have read the journal's policy and the authors of this manuscript have the following competing interests:* **If no** authors have any competing interests to declare, please enter this statement in the box: "*The authors have declared that no competing interests exist.*" * typeset	The authors have declared that no competing interests exist.
Ethics Statement You must provide an ethics statement if your study involved human participants, specimens or tissue samples, or vertebrate animals, embryos or tissues. All information entered here should **also be included in the Methods section** of your	N/A

manuscript. Please write "N/A" if your study does not require an ethics statement.	
Human Subject Research (involved human participants and/or tissue) All research involving human participants must have been approved by the authors' Institutional Review Board (IRB) or an equivalent committee, and all clinical investigation must have been conducted according to the principles expressed in the Declaration of Helsinki. Informed consent, written or oral, should also have been obtained from the participants. If no consent was given, the reason must be explained (e.g. the data were analyzed anonymously) and reported. The form of consent (written/oral), or reason for lack of consent, should be indicated in the Methods section of your manuscript.	
Please enter the name of the IRB or Ethics Committee that approved this study in the space below. Include the approval number and/or a statement indicating approval of this research.	
Animal Research (involved vertebrate animals, embryos or tissues) All animal work must have been conducted according to relevant national and international guidelines. If your study involved non-human primates, you must provide details regarding animal welfare and steps taken to ameliorate suffering; this is in accordance with the recommendations of the Weatherall report, "The use of non-human primates in research." The relevant guidelines followed and the committee that approved the study should be identified in the ethics statement.	
If anesthesia, euthanasia or any kind of animal sacrifice is part of the study, please include briefly in your statement which substances and/or methods were applied.	
Please enter the name of your Institutional Animal Care and Use Committee (IACUC)	

or other relevant ethics board, and indicate whether they approved this research or granted a formal waiver of ethical approval. Also include an approval number if one was obtained. **Field Permit** Please indicate the name of the institution or the relevant body that granted permission.	
Data Availability PLOS journals require authors to make all data underlying the findings described in their manuscript fully available, without restriction and from the time of publication, with only rare exceptions to address legal and ethical concerns (see the PLOS Data Policy and FAQ for further details). When submitting a manuscript, authors must provide a Data Availability Statement that describes where the data underlying their manuscript can be found. Your answers to the following constitute your statement about data availability and will be included with the article in the event of publication. **Please note that simply stating 'data available on request from the author' is not acceptable.** ***If, however, your data are only available upon request from the author(s), you must answer "No" to the first question below, and explain your exceptional situation in the text box provided.*** Do the authors confirm that all data underlying the findings described in their manuscript are fully available without restriction?	No - some restrictions will apply
Please describe where your data may be found, writing in full sentences. **Your answers should be entered into the box below and will be published in the form you provide them, if your manuscript is accepted.** If you are copying our sample text below, please ensure you replace any instances of XXX with the appropriate details. If your data are all contained within the paper and/or Supporting Information files, please state this in your answer below. For example, "All relevant data are within the paper and its Supporting Information files."	Underlying data (Farm Business Survey for England & Wales) analysed are available from the UKDataService: https://discover.ukdataservice.ac.uk/series/?sn=200018.

If your data are held or will be held in a public repository, include URLs, accession numbers or DOIs. For example, “All XXX files are available from the XXX database (accession number(s) XXX, XXX)." If this information will only be available after acceptance, please indicate this by ticking the box below. If neither of these applies but you are able to provide details of access elsewhere, with or without limitations, please do so in the box below. For example: “Data are available from the XXX Institutional Data Access / Ethics Committee for researchers who meet the criteria for access to confidential data.” “Data are from the XXX study whose authors may be contacted at XXX.” * typeset	
Additional data availability information:	

COLEG Y GWYDDORAU NATURIOL
COLLEGE OF NATURAL SCIENCES

YSGOL YR AMGYLCHEDD, ADNODDAU NATURIOL A DAEARYDDIAETH
SCHOOL OF ENVIRONMENT, NATURAL RESOURCES AND GEOGRAPHY

PRIFYSGOL
BANGOR
UNIVERSITY

March 8, 2018

2nd Revision of PONE-D-17-37318

Dear Dr Georgantzis,

Please find enclosed the revised manuscript of entitled "*Metrics and methods for characterizing dairy farm intensification using farm survey data*" for consideration for publication (note change of title).

As requested we have prepared a point-by-point response but highlight the main changes here

Minor edits & changes. Both reviewers suggested some very helpful edits and we have done all of these. In addition, we have taken the opportunity to read through the paper again and made some additional minor changes

Additional statistical explanation. Reviewer #2 while recognising that the manuscript has been improved asks for further statistical explanation. We further explained the statistical approach. The main addition is Figure 1 which shows the workflow of the statistical analysis. This was a great suggestion and really improves the manuscript. We have also added further written explanation to the Data Analysis section. Reviewer #2 also asked that we illustrate the methods with a simplified data set, for space and emphasis reasons we have not done this within the manuscript but point to a book with a freely available PDF that illustrates the approaches with simple data sets.

I look forward to hearing from you.
Yours sincerely,

James Gibbons

PRIFYSGOL BANGOR
BANGOR, GWYNEDD,
LL57 2UW, DU

FFÔN: +44 (01248) 382281
FFACS: +44 (01248) 354997

BANGOR UNIVERSITY
BANGOR, GWYNEDD,
LL57 2UW, UK

TEL: +44 (01248) 382281
FAX: +44 (01248) 354997

M.A. McDONALD, BSc, PhD
PENNAETH YR YSGOL / HEAD OF SCHOOL

RHIF UNIONGYRCHOL / DIRECT LINE: +44 (01248) 388076
EBOST/EMAIL: M.MCDONALD@BANGOR.AC.UK

WWW.BANGOR.AC.UK WWW.BANGOR.AC.UK/SENRGY

Metrics and methods for characterizing dairy farm intensification using farm survey data

Alejandra Gonzalez-Mejia[1], David Styles[1], Paul Wilson[2], James Gibbons[1*]

[1]SENRGy, Bangor University, Deiniol Road, Bangor, LL57 2UW

[2]School of Biosciences, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, LE12 5RD

[*]Corresponding author

Email: j.gibbons@bangor.ac.uk

ABSTRACT

Evaluation of agricultural intensification requires comprehensive analysis of trends in farm performance across physical and socio-economic aspects, which may diverge across farm types. Typical reporting of economic indicators at sectorial or the “average farm” level does not represent farm diversity and provides limited insight into the sustainability of specific intensification pathways. Using farm business data from a total of 7281 farm survey observations of English and Welsh dairy farms over a 14-year period we calculate a time series of 16 key performance indicators (KPIs) pertinent to farm structure, environmental and socio-economic aspects of sustainability. We then apply principle component analysis and model-based clustering analysis to identify statistically the number of distinct dairy farm typologies for each year of study, and link these clusters through time using multidimensional scaling. Between 2001 and 2014, dairy farms have largely consolidated and specialized into two distinct clusters: more extensive farms relying predominantly on grass, with lower milk yields but higher labour intensity, and more intensive farms producing more milk per cow with more concentrate and more maize, but lower labour intensity. There is some indication that these clusters are converging as the extensive cluster is intensifying slightly faster than the intensive cluster, in terms of milk yield per cow and use of concentrate feed. In 2014, annual milk yields were 6,835 and 7,500 l/cow for extensive and intensive farm types, respectively, whilst annual concentrate feed use was 1.3 and 1.5 tonnes per cow. For several KPIs such as milk yield the mean trend across all farms differed substantially from the extensive and intensive typologies mean. The indicators and analysis methodology developed allows identification of distinct farm types and industry trends using readily available survey data. The identified groups allow the accurate evaluation of the consequences of the reduction in dairy farm numbers and intensification at national and international scales.

KEY WORDS

Model-based clustering, PCA, farm typology, intensification, dairy systems, key performance indicators, trend analysis

INTRODUCTION

Globally, dairy production emits 2,128 Mt CO_2e yr^{-1} (roughly 5% of global anthropogenic emissions) and is responsible for a large share of environmental burdens including nutrient losses to air and water, water consumption and land use (1). Demand for dairy products is rising which will lead to a further increase in burdens unless production efficiency increases. One route to this is to reduce land-use intensity of milk production by increasing milk yields per cow (2). However, without advances in technology an environmental gain will only be achieved if the increase in production per cow out paces the increase in demand.

Despite already high milk yields per cow observed in many industrialised countries such as the United Kingdom (UK), dairy production continues on a long-term trend of reduction in farm numbers (consolidation) and intensification (C&I) that is driven by socio-economic and policy factors (3). The UK dairy industry is the 10th largest global producer of cow milk (accounting for 2.2% of world production) (4) and an exemplar of worldwide intensification trends. Between 2001 and 2014, the number of dairy farms in England and Wales decreased by 49%, from 20,191 to 10,274 (5), and the number of dairy cows decreased by 18%, whereas the average number of dairy cows per holding increased by 54%, from 87 to 134 (6), and the average annual milk yield (litres/cow) increased from 6,346 to 7,897 (7). In other words, many farms have exited the sector, whilst remaining farms have grown in size and implemented more intensive practices that support higher milk yields. This trend is expected to continue following the abolition of milk quotas in 2015. However, there is little published information on changes in management and key performance indicators (KPIs) across individual farms, or types of farms, associated with this trend (8). Alvarez et al. (9) emphasize the importance of finding the relationship

56 between intensification and efficiency of dairy farming, and note the lack of studies researching dairy
57 farm heterogeneity hidden behind sectoral statistics.

58 There is high variance in apparent dairy farm management efficiency, as indicated by KPIs such as
59 nutrient use efficiency (10) and grass utilisation efficiency (the proportion of grass grown that is used by
60 dairy cows (11)). Given this range in efficiency it might be expected that intensification of dairy farms
61 will result in more efficient farms growing at the expense of less efficient farms. However, expanding
62 French beef farms are becoming less economically efficient (12) because, despite an increase in
63 investment in capital, technology and concentrate feed output has remained constant since 1990.

64 There are multiple measures of intensification such as the increase in farm output, herd size, feed
65 concentrate use per unit of land or per head, produce per head and produce per unit of land (13).
66 Individually these indicators do not capture all dimensions of farm intensification and do not reflect the
67 sustainability of that intensification (14). Previous studies have assessed aspects of intensification and
68 sustainability (15,16) through the application of productive efficiency methods such Stochastic Frontier
69 Analysis (17,18) or the non-parametric method Data Envelopment Analysis (19). There remains a need
70 to characterise farm intensification beyond these economic and technical efficiency metrics in order to
71 evaluate sustainable intensification.

72 One suggestion(20) is representing dairy systems with multiple derived variables that can be evaluated
73 through the application of Principal Component Analysis (PCA) and clustering analysis. Clustering
74 analysis has previously been applied to i) investigate whether intensification could improve the
75 economic efficiency of dairy farms (9), ii) to classify dairy systems and compare them in terms of
76 productivity, milk destination, maintenance of livestock biodiversity, land management, and landscape
77 conservation (21), and iii) to explore social aspects such as factors that are relevant to quality of life for
78 family dairy farms (22). Here we build on these previous PCA and clustering approaches, using more

robust statistical methods, to define dairy farm typologies according to wider socio-economic characteristics and physical parameters that can be linked to environmental performance and the derivation of carbon, land and nutrient footprints and potentially wider indirect (global) impacts.

We employ KPIs derived from detailed farm survey data to characterize dairy farm production and C&I. Consolidation is measured by the annual reduction in UK dairy farm numbers, and the sustainability of intensification is assessed in terms of physical and socio-economic characteristics critical to environmental, social and economic dimensions of sustainability, including: land use (e.g. grass and fodder) and tenure (i.e. owner occupied area), concentrate feed use, labour intensity, herd size and densities, productivity (i.e. milk yield), and milk price premium received.

METHODS

We used all available England & Wales Farm Business Survey (FBS) dairy farm data, providing 728 dairy farms in 2001 (out of a total all-farm survey population of 2845 across all farm types) declining to 432 farms in 2014 (out of a total all-farm survey population of 2447). These data are available under special license from the UK Data Archive (23–36). Based on KPIs we identified major typologies of farms based on PCA and Clustering Analysis and then investigated how these KPIs and typologies have changed over a 14-year period characterized by structural change. We restricted our sample to farms that had on average at least 10 dairy cows in a calendar year. We then examined relationships among KPIs to identify groups of KPIs that measure particular aspects of farm structure. We also assessed whether relationships among KPIs remain constant over time i.e. if relationships were influenced by structural change (significant differences). The sample was then classified with a model-based clustering method that identified cohorts of similar dairy farms. We then examined changes in these cohorts (clusters) over the study period, 2001 to 2014.

Farm Survey data

Data representing physical-environmental and socio-economic characteristics of dairy farm businesses in England and Wales were extracted from the annual FBS, UK feed (37) and milk prices (38) from 2001 to 2014. Forty-eight variables were extracted annually to calculate 16 KPIs from 7281 farm business observations over 14 years of study. A total of 349,488 data points were analysed. The sample number of cows accounted for in the annual FBS data represents 4-5% of the dairy cow population in Wales and England (2001-2014). See Table S1 for summary of farms included.

The FBS was selected as a data source because it is a comprehensive source of information on socio-economic and physical characteristics of farms including labour, crops (previous and current harvest year, set-aside, by-products, forage and cultivations), livestock (cattle, dairy and other), costs (variable and fixed), assets, enterprise outputs, margins, and incomes. This authoritative source of information is based on a uniform sampling rate that ensures adequate coverage for analysis. Over the sample period farms remained in the survey for up to 15 years, with a replenishing rate of roughly 10% (39).

Key Performance Indicators

We developed an approach to characterize dairy farms based on physical characteristics and production parameters that can be easily derived from farm survey data (Table 1). Our farm characterization is based on widely used variables and indicators that have been applied to represent the structure of dairy farming, its efficiency and the effects of C&I in the dairy business. We developed a set of KPIs using the underlying FBS survey data, but maximised information by transforming descriptors into quantities directly related to measures of production intensity, efficiency and other farm characteristics. We largely excluded economic parameters related to input and output prices, which are exogenous to the farms, but did include a measure of relative price received for milk (an indicator of a milk price premium). The KPIs were derived from widely used indicators to evaluate performance i.e. herd size,

124 stocking rate, herd replacement rate, milk yield, feed amount or cost per animal, and labour
125 requirements (40–42). We added additional indicators such as areas of grass, fodder and cash crops to
126 provide information on land use and feeding strategies that can be used to characterise farms. The
127 agricultural area was also divided into two main areas; one utilised exclusively to grow and harvest crops
128 for human consumption namely, "cash crop", and the "non-cash crop" that is mainly for animal
129 maintenance and that includes fallow, permanent and temporal grass (hay, silage, and grazing including
130 rough grazing), silage cereals, and fodder crops (e.g. roots, kale, and maize) areas. The selected KPIs
131 represent important characteristics of dairy farms with respect to sustainability and intensification,
132 whilst avoiding duplication of information. To give equal weight during the statistical analysis, the KPIs
133 were scaled by the annual mean value for each parameter but results are back scaled and presented in
134 the original KPI units.

135 *Data analysis*

136 We use a suite of statistical methods and workflow to analyse the data as shown in Fig 1. Further details
137 of all the analysis methods with illustrations on simple data sets are available in (43,44) in particular we
138 recommend chapter 14 of Hastie et al. All code to reproduce the data analysis is available on request
139 from the authors. PCA (principal components analysis) was used to explore the relationship among KPIs
140 (i.e. identification of fundamental farm properties) and how these relationships change over time. The
141 usual aim is to reduce multiple dimensions down to two or three for illustration and analysis purposes.
142 PCA creates new linear combinations of existing variables (components) ranked to explain as much
143 variation as possible. The relative weighting of each KPI on each component is then termed the loading
144 and value each farm on the component the score. For the set of KPIs to be a useful measure of farms
145 over time, the relationship between KPIs should be relatively constant but change should result in farms
146 changing their position along the KPI dimensions. PCA was calculated in R (45) and Procrustes rotation of

147 the first 3 KPI loadings identified by PCA was used to compare the structure of each year and compare
148 structure between years with the *vegan* package (46). The Procrustes analysis rotated the PCA loadings
149 to minimize the sum of squares of the difference in distance between loading for each year pair, a small
150 total sum of squares indicating the relationship between the individual KPIs between years was similar, a
151 large difference that the relationship changed between years. The rotation is necessary to fairly
152 compare between years as the relationship between the variables and hence the relative loadings may
153 remain constant over time but the absolute loadings may change and the sign of component loadings is
154 arbitrary (can be positive or negative depending on the algorithm or data used).

155

156 **Fig 1- Statistical workflow used to analyse the Key Performance Indicators (KPIs). Number of clusters**
157 **selected was determined by BIC (Bayesian Information Criterion).**

158

159 Farms were clustered using Gaussian mixture model-based clustering with the *mclust* package in R
160 (47,48). In this method data are considered to originate from a distribution that is a combination of two
161 or more components (i.e. clusters). Each component is modelled by a Gaussian distribution that is
162 characterized by a mean vector, a covariance matrix, and an associated probability in the mixture. Each
163 data point has 16 dimensions (KPI values) with a probability of belonging to each cluster. The model
164 parameters are estimated using the Expectation Maximization algorithm initialized by hierarchical
165 model-based clustering. Each cluster is centred at the mean with increased density for points near the
166 mean (49).

167 We selected this method because the traditional clustering methods (k-means etc.) are heuristic and are
168 not based on formal models with little statistical guidance on number of clusters. Further, the implicit
169 assumptions that clusters are spherical and of equal size are very restrictive when, for example, we

170 might expect there to be small cluster for rarer farm types and larger cluster for common farm types.
171 Trials of k-means and k-medoids clustering on farm survey data performed poorly with a very unstable
172 number of clusters identified. Another advantage of the model-based method is the flexibility of
173 selection for the groups made by geometric features (shape, volume, orientation) of each cluster, which
174 are determined by the covariance matrix. Different model options in *mclust* package are represented by
175 identifiers e.g.: EVI, VEV and VVV. The first identifier denotes volume (equal or variable size), the second
176 shape (spherical or not) and the third orientation (aligned or not). Accordingly, E stands for "equal", V
177 for "variable" and I for "coordinate axes". For example, EVI denotes a model in which the volumes of all
178 clusters are equal (E), the shapes of the clusters may vary (V), and the orientation is the identity (I) or
179 coordinate axes. If all clusters were EEE the results would be similar to k-means clustering. Maximum
180 likelihood is used to fit all these models, with different covariance matrix parameterizations, for a range
181 of components. The best model was selected using the Bayesian Information Criterion or BIC; a small BIC
182 score indicates strong evidence for the corresponding model (47). BIC here trades off degree of model
183 fit against model complexity. Model complexity increases with number of clusters and varying shape,
184 orientation and volume of each cluster.

185 As the clustering was performed independently by year we then used multidimensional scaling (MDS) in
186 order to group similar clusters based on their mean values for each KPI over time and track temporal
187 changes of the same group. We tested the number of dimensions required to well-represent the clusters
188 in ordination space. In this space, clusters more similar in their mean KPI values were closer in terms of
189 ordination distance. For display we ranked within-year clusters by milk yield within year, which means
190 that e.g. cluster 1 in 2001 does not necessarily correspond to cluster 1 in 2002.

Table 1 – Key performance indicators derived from FBS statistics in order to compare the intensity of production and characteristics among farms.

Farm metric		Units	Formula and description	Application
Milk Production	Total dairy cows	qty	Number of dairy cows	Herd size comparison
	Milk yield	l/ qty	Milk production / Dairy Cows	Measure of production efficiency. Higher yield generally means less inputs per production unit
	Milk premium	£/l／£/l	Milk Product Revenue / (Milk Products Sold *Average Milk Price)	Milk price received by farm compared to other farms. Premium >1 is desirable and <1 non-desirable
	Concentrate fed	tonne/ LU	Concentrate Feed Cost / (Concentrate Price * animals in Livestock Units (LU))	Feed bought into the farm that embodies upstream land and environmental impact (e.g. resource depletion, GHG emissions) per livestock unit
	Fodder fed	tonne/ LU	Coarse Fodder Cost / (Fodder Price * animals in Livestock Units (LU))	Measure of feed bought into the farm that embodies upstream land and environmental impacts (e.g. resource depletion, GHG emissions) per livestock unit
Intensity of Livestock Production	Cow fraction	qty/ LU	Dairy Cows / All animals in Livestock Units (LU)	Indicates the degree of the specialization and heterogeneity of the livestock enterprise.
	Cow stocking rate	LU/ ha	Cattle in Livestock Units (LU) / Non-Cash Crop Area	Measure of overall farm land use intensity. Useful for characterising farms and comparing management practices
	Livestock density	qty/ ha	Dairy Cows / Non-Cash Crop Area	Measure of land use intensity for dairy cows
	Labour intensity	hours/ ha	Annual worked hours / Farm Area	Indirect measure of technology. Useful for comparing farm productivity, and for socio-economic characterisation
Grass, Fodder and Maize mix	Fodder area	ha/ ha	Fodder Area /Grass Area	Measure of the reliance on fodder in feeding strategy. Could be used for inferring indoor/outdoor systems and land use footprints.
	Grass area	ha/ ha	Maize Area/Grass Area	Measure of maize dependence in feeding strategy. Could be used to infer land use footprints.
	Non-cash crop area in agricultural area	ha/ ha	Non-Cash Crop Area / Agricultural Area	Measure of farm livestock specialisation

Farm Structure for Animals	Grass area in agricultural area	ha/ ha	Grass Area / Agricultural Area	Measure of grass dependence in feeding strategy. Could be used for inferring indoor/outdoor systems. Useful for comparing farm land use footprints
Production Area	Production area	ha/ ha	Agricultural Area / Farm Area	Measures proportion of farm used for agricultural production.
Tenure	Tenure	ha/ ha	Owner Occupied Area / Agricultural Area	Measure of ownership structure and socio-economic characterisation.
Replacement Rate	Heifers	qty/ qty	Heifers / Dairy Cows	Measure of non-productive herd

RESULTS

Relationships among KPIs

The extracted time series from the FBS were used to compute KPIs that describe dairy farms in a 16-dimensional system (see Figure S1 for distribution). Annual PCAs were computed as well as a calculation that includes all data from 2001 to 2014 (Fig 2). Three dimensions of the PCA (PC1, PC2, and PC3) including all data sets from 2001 to 2014 explain approximately 50% of variation (Figure S2). The loadings on the first 3 components broadly represent seven groups of KPIs (correlated in at least two components): i) milk production specifically (dairy cows, milk yield, concentrate feed per LU, and milk premium), ii) intensity and specialisation of livestock production (dairy stocking density, livestock density, dairy fraction, labour, and fodder per LU), iii) grazing prevalence (cash crop and grass presence), iv) grass/forage maize mix, v) production area, vi) tenure, and vii) replacement rate.

Fig 2 – PCA results for all Key Performance Indicator values across all years (2001-2014). Panels on the left show the PCA scores for individual farms, on the right loading for individual metrics.

Area of land tenured by the owner of a farm is inversely related to dairy production area and replacement rate, which indicates that more heterogeneous farms with low replacement rates are more likely than more specialised dairy farms to be tenured by their owners (Fig 2).

The component scores in Fig. 2 (left-hand plots) show that the majority of farms are concentrated at the centre of the axes for all years (2001-2014) with some outliers for years before 2006. There is some indication that there is less diversity in farms (points are closer together) in later years.

Procrustes rotation of the first 3 components (Fig 3) illustrates that in the periods 2001-2004 and 2006-2014 there are no large differences in the configuration of annual KPIs (sum of squares close to zero)

while 2005 appears an outlier from all other years. This result suggests that the relationship between KPIs has largely remained stable over time, suggesting that they are reliable measures of farm properties even when structural changes are occurring.

Fig 3 - Procrustes analysis of annual variation in relationships among Key performance Indicators (KPIs) are derived from principle component analysis of annual data over the years 2001 – 2014, based on the sum of squared distances.

Cluster identification

Clustering analysis results indicate the number, configuration, and distinctiveness (mixing probabilities) of clusters for each of the survey years. Different cluster configurations are represented by the model i.e. VVV ellipsoidal, varying volume, shape, and orientation and VEV: ellipsoidal, equal shape. Number of farms decreased in the 14 years of study with the majority of farms distributed in mainly two or three clusters (higher probability). Further, clustering analysis identified three clusters for most years except for 2001 and 2003, which had four clusters, and 2011, 2012, and 2014, which had two clusters (Table 2). The distribution of farms among clusters was fairly even in most years with the exception of the smaller clusters (mixing probability < 0.1) (Table 2 & Fig 3). It is likely that these fluctuations in the smaller clusters are a combination of: (i) sampling artefacts where relatively rare farm configurations drop in and out of the sample; (ii) farms that are in transition, or: (iii) farms that have been affected by extreme events. In the majority of years, the individual clusters varied in volume, shape and orientation (VVV) although in a few years (2007, 2009, & 2010) clusters had equal shape (VEV) (Table 2).

Table 2 – Clustering analysis results, indicating the number, configuration and distinctiveness (mixing probabilities) of clusters for each of the survey years.

Year	Cluster configuration	Number of clusters	log likelihood	n	df	Mixing probabilities			
						1	2	3	4
2001	VVV	4	1611	724	611	0.22	0.23	0.35	0.20
2002	VVV	3	431	678	458	0.50	0.48	0.02	
2003	VVV	4	862	643	611	0.38	0.30	0.30	0.02
2004	VVV	3	-182	512	428	0.48	0.37	0.16	
2005	VVV	3	32	477	458	0.42	0.52	0.06	
2006	VVV	3	393	464	458	0.42	0.35	0.23	
2007	VEV	3	67	469	428	0.46	0.42	0.12	
2008	VVV	3	337	493	458	0.55	0.42	0.03	
2009	VEV	3	366	488	428	0.47	0.44	0.09	
2010	VEV	3	623	479	428	0.40	0.15	0.45	
2011	VVV	2	390	479	305	0.37	0.63		
2012	VVV	2	454	467	305	0.44	0.56		
2013	VVV	3	1122	455	458	0.48	0.39	0.12	
2014	VVV	2	505	432	305	0.56	0.44		

Because the clustering analysis was performed for each year separately (a total of 41 clusters in 14 years of study, Table 2), we computed a MDS across clusters from all years in order to track the same type of cohort from one year to another. MDS in a single dimension had an almost perfect linear fit (R=0.999) between the ordination distance and the observed dissimilarity. Thus, this second classification allows changes to be tracked over time for the same type of cluster, and the comparison of clusters. The results show two predominant clusters through time (groups 1 & 2) and other smaller groups (Fig 4). Since there are two main clusters from 2001 to 2014 and four more clusters that appear infrequently, we focused the following discussion on the two predominant clusters: group 1 (circles), classified as "*extensive systems";* group 2 (triangles), categorized as "*intensive systems"* (Fig 4). Note that the outlying groups could comprise very intensive or "mega dairy" farms (Fig 4). All clusters are coloured by their MDS value; so similar clusters can be tracked over time and separated from less common typologies (Fig 4).

Fig 4 – Trends in mean Key Performance Indicator values for all identified clusters over the period 2001 – 2014. The number of farms in each cluster is represented by the size of symbol. Intensive systems are represented by triangles and extensive systems by circles. The solid black line represents the KPI annual average. The distance among all clusters in all years of study is represented by the colour scale MDS. This distance allows identifying which clusters are more similar.

Cluster comparison and trends

KPIs such as number of dairy cows, milk yield, concentrate use, grass in agricultural area, as well as productive area have increased in the 14 years of study (Fig 4); while the other KPIs have remained fairly constant, except for labour intensity, where lower labour use per hectare has been observed over time. Note that the black line that represents the KPI annual average at sectoral level rarely explains the actual value of a cluster for a particular year; extensive and intensive systems are generally above or below.

Fig 4b shows the difference in milk production between clusters. The intensive farms consistently produced more milk per cow than extensive systems. In 2014, there was an annual difference of 665 l/dairy cow in productivity and 8 hours/ha in labour intensity (Fig 4j). Intensive systems had an additional 40 dairy cows (Fig 3a) that consume 224 kg/yr more concentrate per LU (Fig 4c), 21 kg/yr less coarse fodder per LU (Fig 4i), and had a higher ratio of fodder and maize to grass (Fig 4n-o) (Table 3). There was no consistent difference in milk price premiums between intensive and extensive systems over the study period (Fig 4d), though intensive systems fared slightly better for more years.

Despite the difference in milk production, the replacement rate (Table 3), inferred from heifer to cow ratio (Fig 3e), has remained similar for intensive and extensive farms over the 14 years under study, initially declining for both clusters before increasing again towards a peak in 2012. Intensive farms have higher overall stocking densities (Fig 4g-h), but not necessarily a higher dairy fraction (3f) than extensive farms, though differences are small. In 2014, both systems had 1.4 dairy cows per hectare and 2 LU/ha (Table 3). Intensive and extensive systems can also be differentiated by the utilization rate of non-cash crop area (Fig 4l) and grass (Fig 4m) in the agricultural area of a farm, which have not changed dramatically since 2001 for each system (Table 3). Extensive systems utilised almost all non-cash crop agricultural area for grass production, compared to more intensive systems that only used 70% for grass production (Fig 4l-m and Table 3). Intensive farms produced maize on an area equivalent to 20% of grass area, and included fodder areas that grew over time, whilst extensive farms did not produce maize and are characterised by very small fodder areas compared to their grass extent (Fig 4o). There are a small number of intensive farms represented in the sporadic small clusters that appear in some years with comparatively very large maize areas (Fig 4o). Between 2001 and 2014, productivity (l cow^{-1}) in extensive systems increased by 17% and cow numbers by 52%, compared with increases of 13% and 17% in productivity per cow and cow numbers, respectively, for intensive systems (Table 3). At the same time, labour intensity (hours/ha) has declined by 17-18% and the use of concentrate feed has increased

292 by 500-600 kg/LU/yr in both intensive and extensive systems over the study period (Table 3). Across all
293 farm types there was a large increase in the use of concentrates between 2005 and 2008 (Fig 3c).

294 Extensive systems were characterised by a 10% lower rate of owner occupation in 2001, which
295 converged to a similar rate as for intensive systems in 2014, at around 60% of tenure (Fig 4k) (Table 3).
296 Non-agricultural area, such as woodland, buildings, roads, water, and household gardens account for
297 just 3% of farm areas across both intensive and extensive systems (Table 3).

298

Table 3 – Comparison of Key Performance Indicators (KPIs) between 2001 and 2014 for extensive (E) and intensive (I) farm cluster.

KPIs	E 2001	E 2014	I 2001	I 2014	E[2014 -2001]	I[2014-2001]	2001 [E-I]	2014 [E-I]
Total dairy cows	78	132	145	172	55	27	68	40
Milk yield	5,784	6,835	6,588	7,499	1,051	911	-804	- 665
Milk premium	1.01	0.98	1.04	0.97	-0.03	-0.07	-0.03	-0.01
Concentrate fed	0.77	1.29	0.91	1.52	0.52	0.61	-0.13	- 0.22
Fodder fed	0.25	0.16	0.25	0.14	-0.09	-0.11	0.00	-0.02
Cow fraction	0.6	0.6	0.6	0.6	0.1	0.0	0.0	0.0
Cow stocking rate	1.9	2.0	2.1	2.1	0.1	0.0	-0.2	-0.1
Livestock density	1.2	1.4	1.4	1.4	0.1	0.0	-0.2	0.0
Labour intensity	82	68	71	60	-14	-11	10	8
Fodder area	0.0	0.0	0.2	0.3	0.0	0.1	-0.2	-0.3
Grass area	0.0	0.0	0.2	0.2	0.0	0.1	-0.2	-0.2
Non-cash crop area in agricultural area	1.0	1.0	0.9	0.8	0.0	0.0	0.1	0.1
Grass area in agricultural area	1.0	1.0	0.7	0.7	0.0	-0.1	0.3	0.3
Production area	0.97	0.97	0.98	0.97	0.00	0.00	0.00	0.00
Tenure	0.6	0.6	0.5	0.6	0.0	0.0	0.1	0.0
Heifers	0.2	0.2	0.2	0.2	0.0	0.0	0.0	0.0

DISCUSSION

Identifying farm typologies and trends

Analysis of all dairy farm FBS data for the years 2001 through to 2014 confirms a trend of dairy consolidation (3), with a 41% decline in the number of dairy farms surveyed over that time period, and an 18% decline in dairy cow numbers - in line with separate statistics showing that the dairy cow population in England & Wales has declined by 19% over the same period. The total area represented by FBS dairy farms declined by just 7%, reflecting an increase in average dairy farm size from 132 ha in 2001

310 to 141 ha in 2014. Intensification is demonstrated by the 13-17% increase in milk production per cow
311 over the same period. While the FBS data are drawn from a broadly representative set of farms
312 recruited from agricultural census data there will inevitably sampling issues leading to the possible
313 omission of rarer farm types such as very large and intensive dairy farms and farms following a more
314 extensive pathway if these businesses choose not to be surveyed.

315 Many of the 16 KPIs evaluated changed significantly over the 14-year study period, reflecting
316 productivity and efficiency improvements linked to diet change and technological advances.
317 Consolidation is reflected in declining numbers of farms over a slightly declining aggregate area of dairy
318 farms. Intensification is reflected in higher milk yields per cow, higher stocking rates and increased rates
319 of fodder and concentrate feeding over time, coupled with a decline in labour intensity from 2001 to
320 2014 that presumably reflects technological improvements (e.g. investment in more efficient and
321 automated milking parlours) and increased herd size linked to higher performing businesses remaining
322 in dairy; this is in line with previous findings that have demonstrated the importance of efficient labour
323 use in dairy (50,51).

324 We employed a methodology to characterise dairy farm typologies in England and Wales, underpinned
325 by transformation of economic metrics reported in the FBS into KPIs that reflect physical and socio-
326 economic characteristics of dairy farms. While similar approaches have been used previously our
327 methodology is particularly robust because it combines multivariate analyses (PCA) with a Gaussian
328 mixture model-based clustering and multidimensional scaling for grouping similar clusters over time.
329 The method provides robust guidance on the number of clusters to choose. We found that the more
330 usual clustering approaches (k-mean and k-medoid) did not perform well on these data which suggests
331 their use may be inappropriate on FBS and similar data, especially when clusters are not expected to be
332 of equal size or shape. This method can be applied to all agricultural sectors in all countries where farm
333 economic statistics are compiled, providing a solid basis for rigorous and comprehensive evaluation of

334 the sustainability of farm typologies and identified intensification pathways. This approach provides a
335 robust basis for modelling the sustainability of pathways of intensification through time. For example,
336 attributional LCA can be applied to determine the environmental footprints of milk production for
337 statistically defined dairy farm typologies within (e.g. (52)) and between (e.g. (53)) years. Consequential
338 LCA may also be applied to evaluate the environmental loading changes that arise, directly and
339 indirectly, when farm typologies evolve in characteristics and predominance over time (54).

340 Clustering analysis of KPIs identified two main dairy farm typologies, representing different levels of
341 intensification and following different but somewhat convergent pathways of intensification between
342 2001 and 2014. Results indicate that the dairy sector was more heterogeneous in the earlier years of the
343 study, comprising three or four distinct clusters of dairy farms, and consolidated into just two main
344 types of farm that predominated since 2011 (except for 2013). Classification of the two predominant
345 farm typologies as “extensive” and “intensive” may appear to be a simplification, but these types are
346 statistically identified and this does concur with a previous industry report (55) that identified two
347 profitable pathways of dairy farm business development: (i) grass-based expansion to maximise margins
348 per litre of milk; (ii) farm intensification to maximise margins per hectare. Intensive farms achieve higher
349 milk yields per cow but use more concentrate feed and maize than extensive farms, which rely more
350 heavily on grass and do not use maize. Smaller groups of farms identified in some years by clustering
351 analysis had some characteristics indicative of very intensive farming methods, and may reflect a subset
352 of emerging “mega-dairy” farms within the intensive cluster. Results also indicate some degree of
353 convergence between the two main farm typologies, owing to a faster pace of intensification (e.g.
354 increasing milk yield per cow) among the extensive farms. Consequently, in 2014, intensive and
355 extensive farms had the same stocking densities per hectare.

356 *Evaluating the sustainability of intensification*

357 Ongoing C&I is shifting UK dairy production away from small and medium sized farms towards larger
358 farms that can be broadly categorised as grazing- or indoor- dominated systems. C&I pathways influence
359 animal diets, health, yields, grassland and manure management, with implications for environmental
360 and economic efficiency at animal, farm and system level. Whilst the definition of sustainable
361 intensification is contested and may have different meaning in different contexts (56), a broad definition
362 is to raise productivity and social welfare while reducing environmental impacts per unit of output. The
363 measures captured through the use of farm survey data only include a small subset of those in a recent
364 meta-analysis (14). A more complete analysis would require socio-economic, biodiversity and soil health
365 indicators. There is some evidence that environmental and economic indicators may be correlated but
366 social indicators differ (50,57). However, any regional or national analysis upscaling from farms requires
367 being able to identify typologies of farms for which these indicators could be collected in a targeted
368 manner. The indicators developed here can be linked to environmental performance, for example, feed
369 strategies and land use that embody upstream land and environmental impact (e.g. land use, resource
370 depletion, GHG emissions). Therefore, results of this research can be used to model scenarios including
371 social aspects (e.g. labour intensity), economic components (e.g. profits per litres of milk), and
372 environmental impacts (e.g. carbon, land and water footprints) of dairy farming. The clusters also
373 provide a more accurate profile of trends in the sector than hitherto provided by analysis of "average"
374 farms or aggregate data. There is also potential for application in terms of farm management as the
375 developed KPIs could also be used to benchmark farms within cluster typologies, for example in terms of
376 feed use efficiency, and to recommend priority practises to sustainably intensify that are targeted to the
377 distinct cluster typologies.

378 We show that the UK dairy sector can be characterised by 2-3 clusters over the period of study, which
379 allows environmental footprints to be readily calculated using LCA methods. Notably, across all clusters

380 concentrate feed use, and by implication the indirect land footprint of dairy production, increased. A
381 comprehensive analysis would require a wider system boundary than the individual farm. Increased
382 maize and concentrate feed has the potential to improve animal-level efficiency and reduce on-farm
383 environmental footprints (58,59), but may not reduce system level footprints owing to possible land use
384 change GHG emissions (58–60). Coupling the evolution of farm typologies described here with feed
385 sourcing statistics and dairy-beef production models would enable a full LCA appraisal of direct and
386 indirect environmental consequences arising from changes to animal diets and beef co-production (54).
387 Findings from this study may be directly transferable to dairy farming in other industrialised countries
388 where similar C&I trends prevail.

389 Thus, this research provides a foundation for further analyses, in particular LCA (52,61) and DEA (62,63)
390 that can address all aspects of sustainable intensification. Such work could represent a significant
391 advance on previous studies of dairy intensification that have primarily focussed on environmental
392 (52,53) or socio-economic factors (9,17–20) in isolation.

393 CONCLUSION

394 Trends in dairy farm intensification are usually reported at the sectoral or "average farm" level,
395 sometimes differentiated into regions or percentage quartiles, in terms of economic outputs, inputs,
396 and margins (64). Although useful for detecting broad trends, this approach does not adequately
397 capture heterogeneity in farm operations, and does not relate business structure to the physical
398 characteristics necessary to fully evaluate the sustainability of intensification trends. In the method
399 developed here indicators calculated from farm business survey data coupled with robust model-based
400 clustering identify the number of groups of farms and trends over time. We show that in England and
401 Wales dairy farms have largely consolidated and specialized into two distinct clusters that now
402 predominate within the sector: one "extensive" cluster of farms relying on expansion of grass-based

milk production, with lower milk yields and labour intensity; one “intensive” cluster of farms producing, on average, more milk per cow with more concentrate and more maize, but fewer hours of labour per hectare. There is some indication that these clusters are converging as the extensive cluster is intensifying slightly faster than the intensive cluster, in terms of milk yield per cow and use of concentrate feed. The statistical characterisation of these groups will allow the accurate evaluation of the consequences of dairy C&I at national and international scales to be advanced.

ACKNOWLEDGEMENTS

The authors acknowledge funding provided by the Sêr Cymru National Research Network for Low Carbon Energy and Environment (NRN-LCEE) and the comments of three anonymous referees.

REFERENCES

1. Steinfeld H, Gerber P, Wassenaar T, Castel V, Rosales M, De Haan C. Livestock’s Long Shadow: Environmental Issues and Options. FAO, Rome, Italy,. 2006.

2. Steinfeld H, Gerber P. Livestock production and the global environment: Consume less or produce better? Proc Natl Acad Sci. 2010;107(43):18237–8.

3. March MD, Haskell MJ, Chagunda MGG, Langford FM, Roberts DJ. Current trends in British dairy management regimens. J Dairy Sci. 2014;97(12):7985–94.

4. AHDB Dairy. Dairy statistics: an insiders guide 2016. 2016.

5. AHDB Dairy. Producer numbers for England and Wales [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/market-information/farming-data/producer-numbers/uk-producer-numbers/

6. AHDB Dairy. UK cow numbers [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/market-

424 information/farming-data/cow-numbers/uk-cow-numbers/

425 7. AHDB Dairy. Average UK Milk Yield [Internet]. 2017. Available from:
426 https://dairy.ahdb.org.uk/market-information/farming-data/milk-yield/average-milk-
427 yield/#.WjpZvxOFiOE

428 8. European Commission. Structure and dynamics of EU farms : changes , trends and policy
429 relevance. EU Agricultural Economics Briefs. 2013;(9):1–15.

430 9. Alvarez A, del Corral J, Solís D, Pérez JA. Does Intensification Improve the Economic Efficiency of
431 Dairy Farms? J Dairy Sci. 2008;91(9):3693–8.

432 10. Van Grinsven HJM, Ten Berge HFM, Dalgaard T, Fraters B, Durand P, Hart A, et al. Management,
433 regulation and environmental impacts of nitrogen fertilization in northwestern Europe under the
434 Nitrates Directive; A benchmark study. Biogeosciences. 2012;9(12):5143–60.

435 11. Dairy Development Centre. Improving the Welsh dairy supply chain: Grass value. Camarthen;
436 2014.

437 12. Veysset P, Lherm M, Roulenc M, Troquier C, Bébin D. Productivity and technical efficiency of
438 suckler beef production systems: Trends for the period 1990 to 2012. Animal. 2015;9(12):2050–9.

439 13. Bava L, Sandrucci A, Zucali M, Guerci M, Tamburini A. How can farming intensification affect the
440 environmental impact of milk production? J Dairy Sci. 2014;97(7):4579–93.

441 14. Mahon N, Crute I, Simmons E, Islam MM. Sustainable intensification – 'oxymoron' or 'third-way'?
442 A systematic review. Ecol Indic. 2017;74:73–97.

443 15. Garnett T, Appleby MC, Balmford A, Bateman IJ, Benton TG, Bloomer P, et al. Sustainable
444 Intensification in Agriculture: Premises and Policies. Science (80-). 2013;341(6141):33–4.

16. Garnett T, Godfray C. Sustainable intensification in agriculture. Navigating a course through competing food system priorities. Food Climate Research Network and the Oxford Martin Programme on the Future of Food. 2012.

17. Cabrera VE, Solís D, del Corral J. Determinants of technical efficiency among dairy farms in Wisconsin. J Dairy Sci. 2010;93(1):387–93.

18. Dong F, Hennessy DA, Jensen HH, Volpe RJ. Technical efficiency, herd size, and exit intentions in U.S. dairy farms. Agric Econ. 2016;47(5):533–45.

19. Toma L, March M, Stott AW, Roberts DJ. Environmental efficiency of alternative dairy systems: A productive efficiency approach. J Dairy Sci. 2013;96(11):7014–31.

20. Brotzman RL, Cook NB, Nordlund K, Bennett TB, Gomez Rivas A, Döpfer D. Cluster analysis of Dairy Herd Improvement data to discover trends in performance characteristics in large Upper Midwest dairy herds. J Dairy Sci. 2015;98(5):3059–70.

21. Sturaro E, Marchiori E, Cocca G, Penasa M, Ramanzin M, Bittante G. Dairy systems in mountainous areas: Farm animal biodiversity, milk production and destination, and land use. Livest Sci. 2013;158(1–3):157–68.

22. Maseda F, Díaz F, Alvarez C. Family dairy farms in Galicia (N.W. Spain): Classification by some family and farm factors relevant to quality of life. Biosyst Eng. 2004;87(4):509–21.

23. Duchy College Rural Business School. Farm Business Survey, 2011-2012: Special Licence Access. [data collection]. 2nd Edition. 2014.

24. Duchy College Rural Business School. Farm Business Survey, 2013-2014: Special Licence Access. [data collection]. 3rd Edition. 2016.

25. Duchy College Rural Business School. Farm Business Survey, 2012-2013: Special Licence Access. [data collection]. 2nd Edition. 2015.

26. Duchy College Rural Business School. Farm Business Survey, 2014-2015: Special Licence Access. [data collection]. 2nd Edition. 2016.

27. Defra, National Assembly for Wales. Farm Business Survey, 2001-2002: Special Licence Access. [data collection]. 2004.

28. Defra, National Assembly for Wales. Farm Business Survey, 2002-2003: Special Licence Access. [data collection]. 2004.

29. Defra, National Assembly for Wales. Farm Business Survey, 2003-2004: Special Licence Access. [data collection]. 2005.

30. Defra, National Assembly for Wales. Farm Business Survey, 2005-2006: Special Licence Access. [data collection]. 3rd Edition. 2008.

31. Defra, National Assembly for Wales. Farm Business Survey, 2006-2007: Special Licence Access. [data collection]. 2008.

32. Defra, National Assembly for Wales. Farm Business Survey, 2007-2008: Special Licence Access. [data collection]. 2010.

33. Defra, National Assembly for Wales. Farm Business Survey, 2008-2009: Special Licence Access. [data collection]. 2nd Edition. 2011.

34. Defra, National Assembly for Wales. Farm Business Survey, 2009-2010: Special Licence Access. [data collection]. 3rd Edition. 2014.

35. Defra, National Assembly for Wales. Farm Business Survey, 2010-2011: Special Licence Access.

[data collection]. 3rd Edition. 2014.

36. Defra, National Assembly for Wales. Farm Business Survey, 2004-2005: Special Licence Access. [data collection]. 2006.

37. AHDB Dairy. Feed prices [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/resources-library/market-information/farm-expenses/feed-prices/#.WBi26-GLRR4

38. Defra. Calendar year farmgate milk prices UK 1970 to 2014 [Internet]. 2016. Available from: https://data.gov.uk/dataset/milk_prices_and_composition_of_milk/resource/42f5e421-b6db-45a4-a5e0-08f65c040413

39. Farm business survey - technical notes and guidance [Internet]. Available from: https://www.gov.uk/guidance/farm-business-survey-technical-notes-and-guidance

40. AHDB Dairy. Evidence Report: GB dairy herd performance 2014/15. Kenilworth; 2015.

41. Redman G. The John Nix Farm Management Pocketbook 2012. 40th ed. Melton Mowbray: The Andersons Centre; 2011. 280 p.

42. SAC Consulting. The Farm Management Handbook. 30th ed. McBain C, Curry J, editors. Edinburgh: SAC Consulting; 2016. 332 p.

43. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Data Mining, Inference, and Prediction. 2nd ed. Springer-Verlag; 2009. 763 p.

44. Mardia KV, Kent JT, Bibby JM. Multivariate Analysis. Academic Press; 1979.

45. R Core Team. R: A language and environment for statistical computing. Vienna, Austria; 2017.

46. Oksanen J, F. Guillaume, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package. R package version 2.0-2. 2011.

508 47. Fraley C, Raftery AE, Murphy TB, Scrucca L. mclust Version 4 for R: Normal Mixture Modeling for
509 Model-Based Clustering, Classification, and Density Estimation. Technical Report 597, University
510 of Washington. 2012.

511 48. Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density estimation. J Am
512 Stat Assoc. 2002;97(458):611–31.

513 49. Fraley C, Raftery AE. Bayesian Regularization for Normal Mixture Estimation and Model-Based
514 Clustering. J Classif. 2007;24(2):155–81.

515 50. Soteriades AD, Faverdin P, March M, Stott AW. Improving efficiency assessments using additive
516 data envelopment analysis models: an application to contrasting dairy farming systems.
517 Agricultural and Food Science. 2015;24(3):235–48.

518 51. Wilson P. Decomposing variation in dairy profitability: The impact of output, inputs, prices,
519 labour and management. J Agric Sci. 2011;149(4):507–17.

520 52. O'Brien D, Shalloo L, Patton J, Buckley F, Grainger C, Wallace M. A life cycle assessment of
521 seasonal grass-based and confinement dairy farms. Agric Syst. 2012;107:33–46.

522 53. Capper JL, Cady RA, Bauman DE. The environmental impact of dairy production: 1944 compared
523 with 2007. J Anim Sci. 2009;87(6):2160–7.

524 54. Styles D, Gonzalez-Mejia A, Moorby J, Foskolos A, Gibbons J. Climate mitigation by dairy
525 intensification depends on intensive use of spared grassland. Global Change Biology. 2017;

526 55. The Andersons Centre. The structure of the GB dairy farming industry – what drives change?
527 2013.

528 56. Loos J, Abson DJ, Chappell MJ, Hanspach J, Mikulcak F, Tichit M, et al. Putting meaning back into

529 ‘sustainable intensification’. Front Ecol Environ. 2014;12(6):356–61.

530 57. Reig-Martínez E, Gómez-Limón JA, Picazo-Tadeo AJ. Ranking farms with a composite indicator of
531 sustainability. Agric Econ. 2011;42(5):561–75.

532 58. Vellinga T V., Hoving IE. Maize silage for dairy cows: Mitigation of methane emissions can be
533 offset by land use change. Nutr Cycl Agroecosystems. 2011;89(3):413–26.

534 59. Van Middelaar CE, Berentsen PBM, Dijkstra J, De Boer IJM. Evaluation of a feeding strategy to
535 reduce greenhouse gas emissions from dairy farming: The level of analysis matters. Agric Syst.
536 2013;121:9–22.

537 60. Styles D, Gibbons J, Williams AP, Stichnothe H, Chadwick DR, Healey JR. Cattle feed or bioenergy?
538 Consequential life cycle assessment of biogas feedstock options on dairy farms. GCB Bioenergy.
539 2015;7(5):1034–49.

540 61. Styles D, Gibbons J, Williams AP, Dauber J, Stichnothe H, Urban B, et al. Consequential life cycle
541 assessment of biogas, biofuel and biomass energy options within an arable crop rotation. GCB
542 Bioenergy. 2015;7(6):1305–20.

543 62. Iribarren D, Hospido A, Moreira MT, Feijoo G. Benchmarking environmental and operational
544 parameters through eco-efficiency criteria for dairy farms. Sci Total Environ. 2011;409(10):1786–
545 98.

546 63. Wilson P, Lewis M, Crane R, Robertson P, Bonner J, Davenport R, et al. Farm Level Performance :
547 Identifying Common Factors Determining Levels of Performance. 2012.

548 64. Mchoul H, Robertson P, Wilson P. Dairy Farming in England. 2011.

549

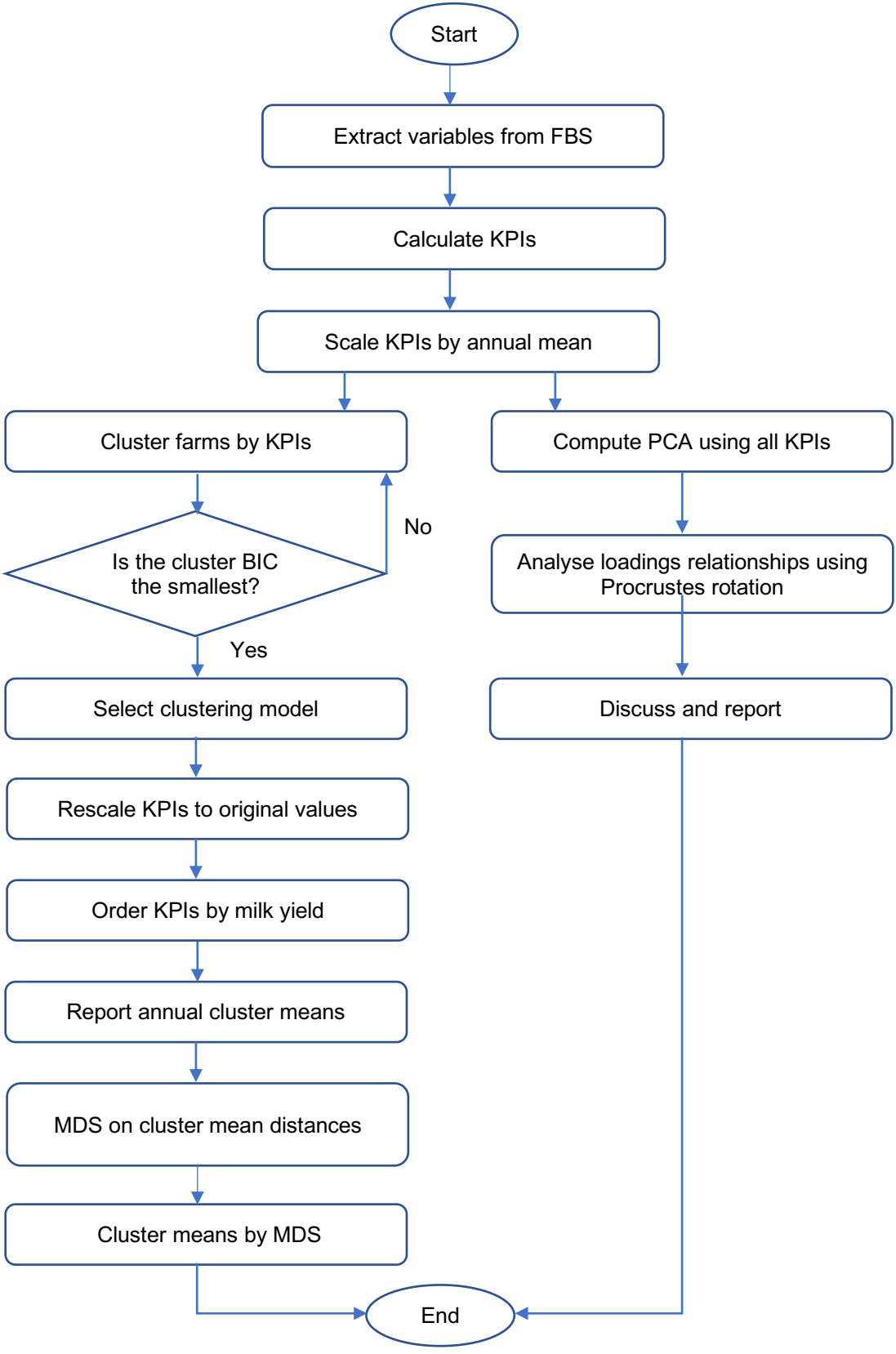
Start
Extract variables from FBS
Calculate KPIs
Scale KPIs by annual mean
Cluster farms by KPIs
Compute PCA using all KPIs
Is the cluster BIC the smallest?
No
Yes
Analyse loadings relationships using Procrustes rotation
Select clustering model
Discuss and report
Rescale KPIs to original values
Order KPIs by milk yield
Report annual cluster means
MDS on cluster mean distances
Cluster means by MDS
End

Figure 2

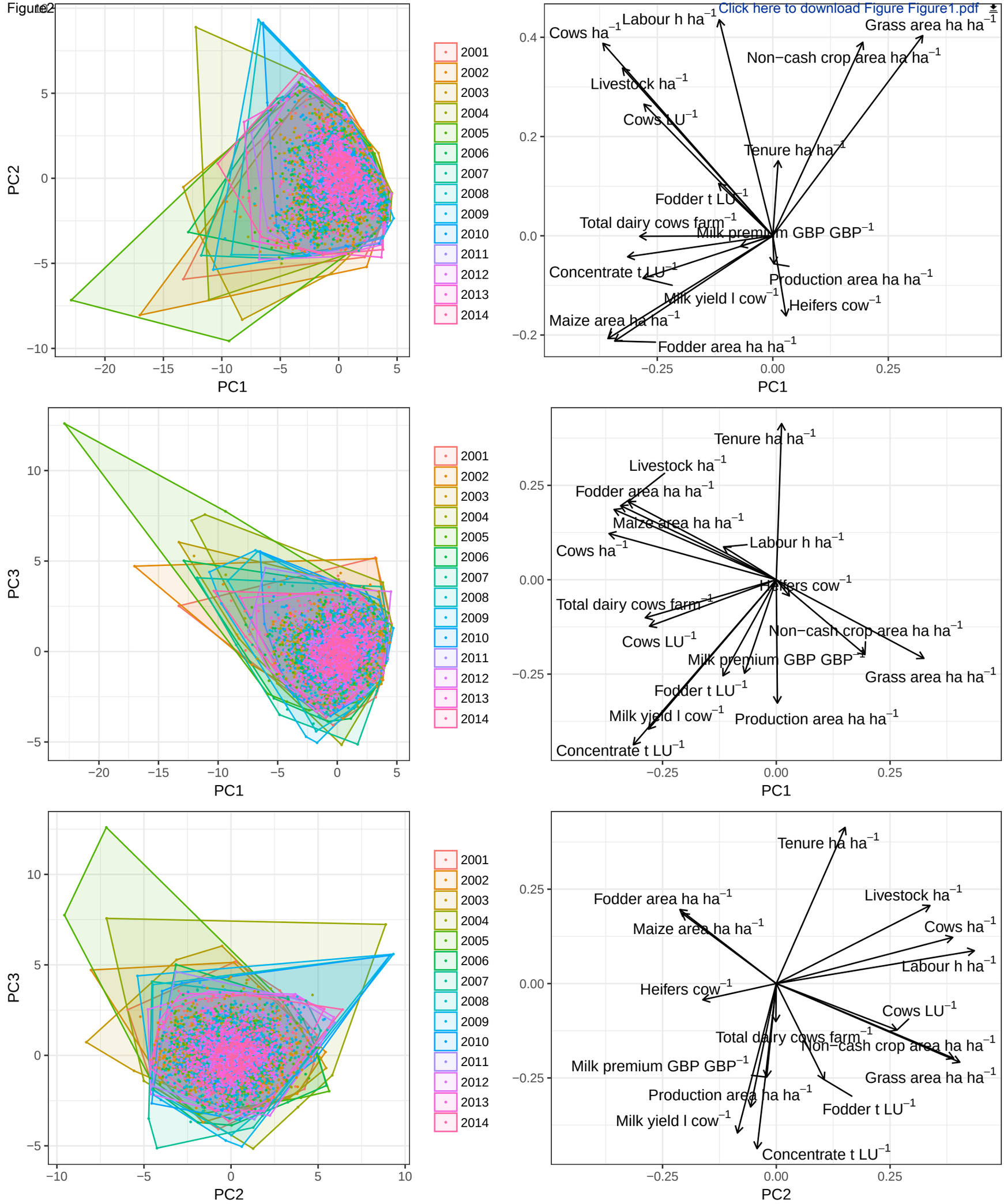

Figure3

Click here to download Figure Procrustes plot years Ale1.pdf

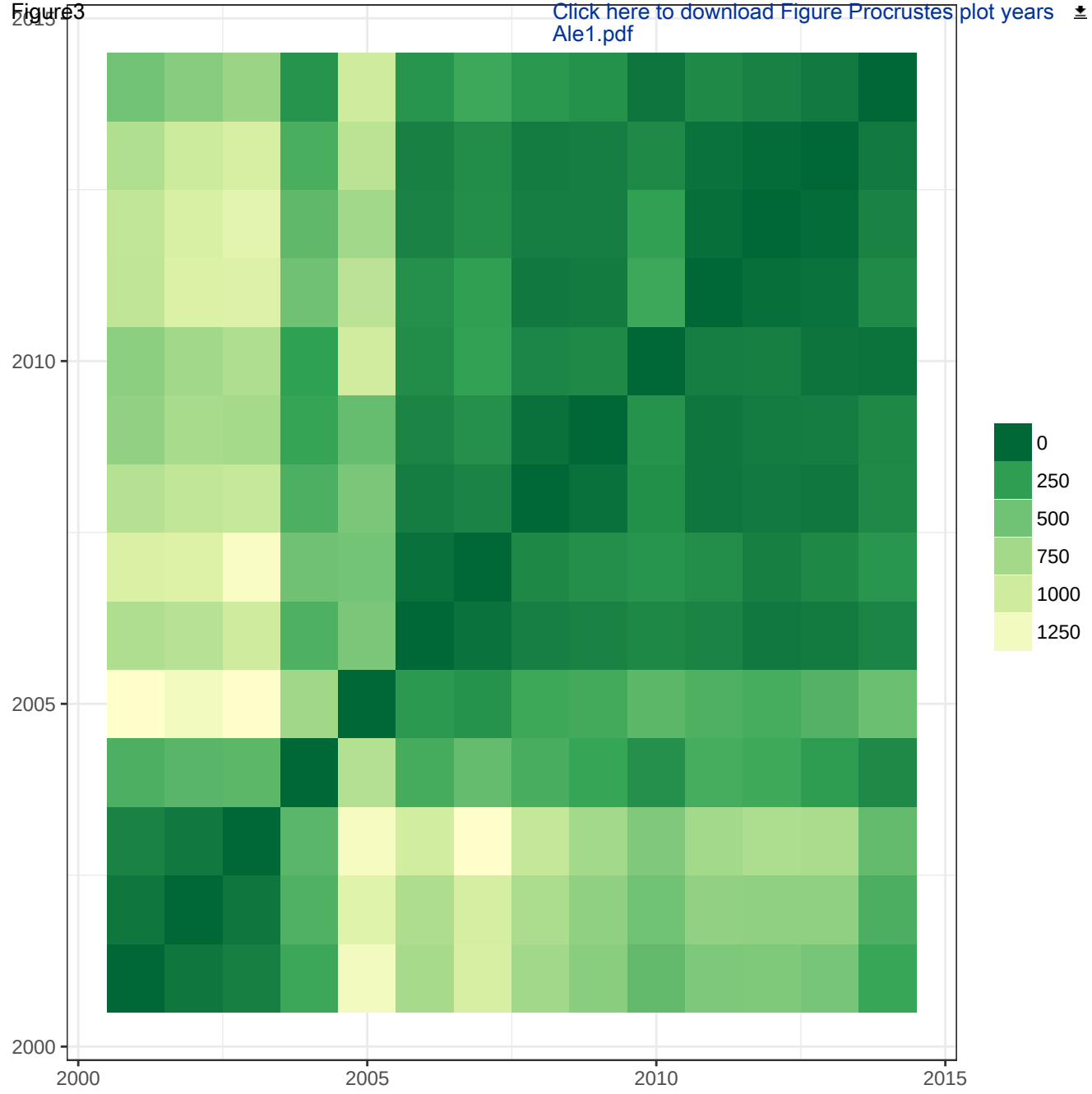

Figure4

Click here to download Figure Figure3.pdf

Metrics and methods for characterizing ~~Dairy Farm Intensification Between 2001 And 2014~~dairy farm intensification using farm survey data

Alejandra Gonzalez-Mejia[1], David Styles[1], Paul Wilson[2], James Gibbons[1*]

[1]SENRGy, Bangor University, Deiniol Road, Bangor, LL57 2UW

[2]School of Biosciences, University of Nottingham, Sutton Bonington Campus, Sutton Bonington, LE12 5RD

[*]Corresponding author

Email: j.gibbons@bangor.ac.uk

ABSTRACT

Evaluation of ~~trends in~~ agricultural intensification requires comprehensive analysis of trends in farm performance across physical and socio-economic aspects, which may diverge across farm types. Typical reporting of economic indicators at sectorial or the "average farm" level ~~may~~does not represent farm diversity and ~~is potentially insufficient to evaluate~~provides limited insight into the sustainability of specific intensification pathways. Using farm business data from a total of 7281 farm survey observations of English and Welsh dairy farms over a 14-year period we calculate a time series of 16 key performance indicators (KPIs) pertinent to farm structure ~~and~~, environmental and socio-economic aspects of sustainability. We then apply principle component analysis and model-based clustering analysis to identify statistically the number of distinct dairy farm typologies for each year of study, and link these clusters through time using multidimensional scaling. Between 2001 and 2014, dairy farms have largely consolidated and specialized into two distinct clusters: more extensive farms relying predominantly on grass, with lower milk yields but higher labour intensity, and more intensive farms producing more milk per cow with more concentrate and more maize, but lower labour intensity. There is some indication that these clusters are converging as the extensive cluster is intensifying slightly faster than the intensive cluster, in terms of milk yield per cow and use of concentrate feed. In 2014, annual milk yields were 6,835 and 7,500 l/cow for extensive and intensive farm types, respectively, whilst annual concentrate feed use was 1.3 and 1.5 tonnes per cow. For several KPIs such as milk yield the mean trend across all farms differed substantially from the extensive and intensive typologies mean. The indicators and analysis methodology developed allows identification of distinct farm types and industry trends using readily available survey data. The identified groups allow the accurate evaluation of the consequences of the reduction in dairy farm numbers and intensification at national and international scales.

KEY WORDS

Model-based clustering, PCA, farm typology, intensification, dairy systems, key performance indicators, trend analysis

INTRODUCTION

Globally, dairy production emits 2,128 Mt CO_2e yr^{-1} (roughly 5% of global anthropogenic emissions) and is responsible for a large share of environmental burdens including nutrient losses to air and water, water consumption and land use (1). Demand for dairy products is rising which will lead to a further increase in burdens unless production efficiency increases. One route to this is ~~for~~to reduce land-use intensity of milk production ~~to be reduced~~ by increasing milk yields per cow (2). However, without ~~other technical improvements~~advances in technology an environmental gain will only be achieved if the increase in production per cow out paces the increase in demand.

Despite already high milk yields per cow observed in many industrialised countries such as the United Kingdom (UK), dairy production continues on a long-term trend of reduction in farm numbers (consolidation) and intensification (C&I) that is driven by socio-economic and policy factors (3). The UK dairy industry ~~is an exemplar of worldwide intensification trends, and~~ is the 10th largest global producer of cow milk (accounting for 2.2% of world production) (4)~~.~~ and an exemplar of worldwide intensification trends. Between 2001 and 2014, the number of dairy farms in England and Wales decreased by 49%, from 20,191 to 10,274 (5), and the number of dairy cows decreased by 18%, whereas the average number of dairy cows per holding increased by 54%, from 87 to 134 (6), and the average annual milk yield (litres/cow) increased from 6,346 to 7,897 (7). In other words, many farms have exited the sector, whilst remaining farms have grown in size and implemented more intensive practices that support higher milk yields. This trend is expected to continue following the abolition of milk quotas in 2015. However, there is little published information on changes in management and key performance

indicators (KPIs) across individual farms, or types of farms, associated with this trend (8). ~~While~~Alvarez et al. (9) emphasize the importance of finding the relationship between intensification and efficiency of dairy farming, and note the lack of studies researching dairy farm heterogeneity hidden behind sectoral statistics.

There is high variance in apparent dairy farm management efficiency, as indicated by KPIs such as nutrient use efficiency (10) and grass utilisation efficiency (the proportion of grass grown that is used by dairy cows (11)). ~~It~~Given this range in efficiency it might be expected that intensification of dairy farms will result in more efficient farms growing at the expense of less efficient farms. However, expanding French beef farms are becoming less economically efficient (12) because, despite an increase in investment in capital, technology and concentrate feed output has remained constant since 1990.

There are multiple measures of intensification such as the increase in farm output, herd size, feed concentrate use per unit of land or per head, produce per head and produce per unit of land (13). Individually these indicators do not capture all dimensions of farm intensification and do not reflect the sustainability of that intensification (14). Previous studies have assessed aspects of intensification and sustainability (15,16) through the application of productive efficiency methods such Stochastic Frontier Analysis (17,18) or the non-parametric method Data Envelopment Analysis (19). There remains a need to characterise farm intensification beyond these economic and technical efficiency metrics in order to evaluate sustainable intensification.

One suggestion(20) is representing dairy systems with multiple derived variables that can be evaluated through the application of Principal Component Analysis (PCA) and clustering analysis. Clustering analysis has previously been applied to i) investigate whether intensification could improve the economic efficiency of dairy farms (9), ii) to classify dairy systems and compare them in terms of productivity, milk destination, maintenance of livestock biodiversity, land management, and landscape

conservation (21), and iii) to explore social aspects such as factors that are relevant to quality of life for family dairy farms (22). Here we build on these previous PCA and clustering approaches, using more robust statistical methods, to define dairy farm typologies according to wider socio-economic characteristics and physical parameters that can be linked to environmental performance and the derivation of carbon, land and nutrient footprints and potentially wider indirect (global) impacts.

We employ KPIs derived from detailed farm survey data to characterize dairy farm production and C&I. Consolidation is measured by the annual reduction in UK dairy farm numbers, and the sustainability of intensification is assessed in terms of physical and socio-economic characteristics critical to environmental, social and economic dimensions of sustainability, including: land use (e.g~~.~~, grass and fodder) and tenure (i.e~~.~~, owner occupied area), concentrate feed use, labour intensity, herd size and densities, productivity (i.e~~.~~, milk yield), and milk price premium received.

METHODS

We used all available England & Wales Farm Business Survey (FBS) dairy farm data, providing 728 dairy farms in 2001 (out of a total all-farm survey population of 2845 across all farm types) ~~in 2001~~ declining to 432 farms in 2014 (out of a total all-farm survey population of 2447~~) in 2014.~~). These data are available under special license from the UK Data Archive (23–36). Based on KPIs we ~~identify~~identified major typologies of farms based on PCA and Clustering Analysis and then ~~investigate~~investigated how these KPIs and typologies have changed over a 14-year period characterized by structural change. We restricted our sample to farms that had on average at least 10 dairy cows in a calendar year. We then examined relationships among KPIs to identify groups of KPIs that measure particular aspects of farm structure. We also assessed whether relationships among KPIs remain constant over time i.e. if relationships were influenced by structural change (significant differences). The sample was then

classified with a model-based clustering method that identified cohorts of similar dairy farms. We then examined changes in these cohorts (clusters) over the study period, 2001 to 2014.

Farm Survey data

Data representing physical-environmental and socio-economic characteristics of dairy farm businesses in England and Wales were extracted from the annual FBS, UK feed (37) and milk prices (38) from 2001 to 2014. Forty-eight variables were extracted annually to calculate 16 KPIs from 7281 farm business observations over 14 years of study. A total of 349,488 data points were analysed. The sample number of cows accounted for in the annual FBS data represents 4-5% of the dairy cow population in Wales and England (2001-2014). See Table S1 for summary of farms included.

The FBS was selected as a data source because it is a comprehensive source of information on socio-economic and physical characteristics of farms including labour, crops (previous and current harvest year, set-aside, by-products, forage and cultivations), livestock (cattle, dairy and other), costs (variable and fixed), assets, enterprise outputs, margins, and incomes. This authoritative source of information is based on a uniform sampling rate that ensures adequate coverage for analysis. Over the sample period farms remained in the survey for up to 15 years, with a replenishing rate of roughly 10% (39).

Key Performance Indicators

We developed an approach to characterize dairy farms based on physical characteristics and production parameters that can be easily derived from farm survey data (Table 1). Our farm characterization is based on widely used variables and indicators that have been applied to represent the structure of dairy farming, its efficiency and the effects of C&I in the dairy business. We developed a set of KPIs using the underlying FBS survey data, but ~~maximising~~maximised information by transforming descriptors into quantities~~, ratios, densities and intensities that~~ directly ~~relate~~related to measures of production

intensity, efficiency and other farm characteristics. We largely excluded economic parameters related to input and output prices, which are exogenous to the farms, but did include a measure of relative price received for milk (an indicator of a milk price premium). The KPIs were derived from widely used indicators to evaluate performance i.e. herd size, stocking rate, herd replacement rate, milk yield, feed amount or cost per animal, and labour requirements (40–42). We added additional indicators such as areas of grass, fodder and cash crops to provide information on land use and feeding strategies that can be used to characterise farms. The agricultural area was also divided into two main areas; one utilised exclusively to grow and harvest crops for human consumption namely, "cash crop", and the "non-cash crop" that is mainly for ~~animals, which~~animal maintenance and that includes fallow, permanent and temporal grass (hay, silage, and grazing including rough grazing), silage cereals, and fodder crops (e.g~~.,~~. roots, kale, and maize) areas. The selected KPIs represent important characteristics of dairy farms with respect to sustainability and intensification, whilst avoiding duplication of information. To give equal weight ~~in~~during the statistical analysis, the KPIs were scaled by the annual mean value for each parameter but results are back scaled and presented in the original KPI units.

Data analysis

~~PCA was used to explore the relationship among KPIs (i.e., identification of fundamental farm properties) and how these relationships change over time. For a set of KPIs to be a useful measure over time, the relationship between KPIs should be relatively constant but change should result in farms changing their position along the KPI dimensions. PCA was calculated in R (43) and Procrustes rotation of the first 3 KPI loadings identified by PCA was used to compare the structure of each year and compare structure between years with the *vegan* package (44). The Procrustes analysis rotated the PCA loadings to minimize the sum of squares of the difference in distance between loading for each year pair, a small total sum of squares indicating the relationship between the individual KPIs between years was similar, a~~

~~large difference that the relationship changed between years. The rotation is necessary as the sign of component loadings is arbitrary.~~

~~Farms were clustered using Gaussian mixture model-based clustering with the *mclust* package in R (45,46). In this method, data are considered to originate from a distribution that is a combination of two or more components (i.e. clusters). Each component is modelled by a Gaussian distribution that is characterized by a mean vector, a covariance matrix, and an associated probability in the mixture. Each data point has 16 dimensions (KPI values) with a probability of belonging to each cluster. The model parameters are estimated using the Expectation Maximization algorithm initialized by hierarchical model-based clustering. Each cluster is centred at the mean with increased density for points near the mean (47).~~

We use a suite of statistical methods and workflow to analyse the data as shown in Fig 1. Further details of all the analysis methods with illustrations on simple data sets are available in (43,44) in particular we recommend chapter 14 of Hastie et al. All code to reproduce the data analysis is available on request from the authors. PCA (principal components analysis) was used to explore the relationship among KPIs (i.e. identification of fundamental farm properties) and how these relationships change over time. The usual aim is to reduce multiple dimensions down to two or three for illustration and analysis purposes. PCA creates new linear combinations of existing variables (components) ranked to explain as much variation as possible. The relative weighting of each KPI on each component is then termed the loading and value each farm on the component the score. For the set of KPIs to be a useful measure of farms over time, the relationship between KPIs should be relatively constant but change should result in farms changing their position along the KPI dimensions. PCA was calculated in R (45) and Procrustes rotation of the first 3 KPI loadings identified by PCA was used to compare the structure of each year and compare structure between years with the *vegan* package (46). The Procrustes analysis rotated the PCA loadings to minimize the sum of squares of the difference in distance between loading for each year pair, a small

total sum of squares indicating the relationship between the individual KPIs between years was similar, a large difference that the relationship changed between years. The rotation is necessary to fairly compare between years as the relationship between the variables and hence the relative loadings may remain constant over time but the absolute loadings may change and the sign of component loadings is arbitrary (can be positive or negative depending on the algorithm or data used).

Fig 1- Statistical workflow used to analyse the Key Performance Indicators (KPIs). Number of clusters selected was determined by BIC (Bayesian Information Criterion).

Farms were clustered using Gaussian mixture model-based clustering with the *mclust* package in R (47,48). In this method data are considered to originate from a distribution that is a combination of two or more components (i.e. clusters). Each component is modelled by a Gaussian distribution that is characterized by a mean vector, a covariance matrix, and an associated probability in the mixture. Each data point has 16 dimensions (KPI values) with a probability of belonging to each cluster. The model parameters are estimated using the Expectation Maximization algorithm initialized by hierarchical model-based clustering. Each cluster is centred at the mean with increased density for points near the mean (49).

We selected this method because the traditional clustering methods (k-means etc.) are heuristic and are not based on formal models~~.~~ with little statistical guidance on number of clusters. Further, the implicit assumptions that clusters are spherical and of equal size ~~is~~are very restrictive when, for example, we might expect there to be small cluster for rarer farm types and larger cluster for common farm types. Trials of k-means and k-medoids clustering on ~~these~~farm survey data performed poorly with a very unstable number of clusters identified. Another advantage of the model-based method is the flexibility

of selection for the groups made by geometric features (shape, volume, orientation) of each cluster, which are determined by the covariance matrix. Different model options in *mclust* package are represented by identifiers e.g.: EVI, VEV and VVV. The first identifier denotes volume (equal or variable size), the second shape (spherical or not) and the third orientation (aligned or not). Accordingly, E stands for "equal", V for "variable" and I for "coordinate axes". For example, EVI denotes a model in which the volumes of all clusters are equal (E), the shapes of the clusters may vary (V), and the orientation is the identity (I) or coordinate axes. If all clusters were EEE the results would be similar to k-means clustering. Maximum likelihood is used to fit all these models, with different covariance matrix parameterizations, for a range of components. The best model was selected using the Bayesian Information Criterion or BIC; a small BIC score indicates strong evidence for the corresponding model ~~(45).~~(47). BIC here trades off degree of model fit against model complexity. Model complexity increases with number of clusters and varying shape, orientation and volume of each cluster.

As the clustering was performed independently by year we then used multidimensional scaling (MDS) in order to group similar clusters based on their mean values for each KPI over time and track temporal changes of the same group. We tested the number of dimensions required to well~~-~~represent the clusters in ordination space. In this space, clusters more similar in their mean KPI values were closer in terms of ordination distance. For display we ranked within~~-~~year clusters by milk yield within year, which means that e.g~~.,~~. cluster 1 in 2001 does not necessarily ~~exactly~~ correspond to cluster 1 in 2002.

Formatted: Line spacing: Double

Table 1 – Key performance indicators derived from FBS statistics in order to compare the intensity of production and characteristics among farms.

Farm metric		Units	Formula and description	Application
Milk Production	~~Dairy Cows~~Total dairy cows	qty	Number of dairy cows	Herd size comparison
	Milk ~~Yield~~yield	l/ qty	Milk production / Dairy Cows	Measure of production efficiency. Higher yield generally means less inputs per production unit
	Milk ~~Price Premium~~premium	£/l ~~/~~ £/l	Milk Product Revenue / (Milk Products Sold *Average Milk Price)	Milk price received by farm compared to other farms. Premium >1 is desirable and <1 non-desirable
	~~Animals~~ Concentrate fed	tonne/ LU	Concentrate Feed Cost / (Concentrate Price * animals in Livestock Units (LU))	Feed bought into the farm that embodies upstream land and environmental impact (e.g~~.,~~. resource depletion, GHG emissions) per livestock unit
	~~Animals Coarse~~ Fodder fed	tonne/ LU	Coarse Fodder Cost / (Fodder Price * animals in Livestock Units (LU))	Measure of feed bought into the farm that embodies upstream land and environmental impacts (e.g. resource depletion, GHG emissions) per livestock unit
Intensity of Livestock Production	~~Dairy Fraction~~Cow fraction	qty/ LU	Dairy Cows / All animals in Livestock Units (LU)	Indicates the degree of the specialization and heterogeneity of the livestock enterprise.
	~~Livestock Density~~Cow stocking rate	LU/ ha	Cattle in Livestock Units (LU) / Non-Cash Crop Area	Measure of overall farm land use intensity. Useful for characterising farms and comparing management practices
	~~Stocking Density~~Livestock density	qty/ ha	Dairy Cows / Non-Cash Crop Area	Measure of land use intensity for dairy cows
	Labour ~~Intensity~~intensity	hours/ ha	Annual worked hours / Farm Area	Indirect measure of technology. Useful for comparing farm productivity, and for socio-economic characterisation
Grass, Fodder and Maize mix	Fodder ~~Grass Ratio~~area	ha/ ha	Fodder Area /Grass Area	Measure of the reliance on fodder in feeding strategy. Could be used for inferring indoor/outdoor systems and land use footprints.
	~~Maize~~ Grass ~~Ratio~~area	ha/ ha	Maize Area/Grass Area	Measure of maize dependence in feeding strategy. Could be used to infer land use footprints.

Farm Structure for Animals	~~Fraction of~~ Non-~~Cash Crop Area~~cash crop area in ~~Agricultural Area~~agricultural area	ha/ ha	Non-Cash Crop Area / Agricultural Area	Measure of farm livestock specialisation
	~~Fraction of~~ Grass area in ~~Agricultural Area~~agricultural area	ha/ ha	Grass Area / Agricultural Area	Measure of grass dependence in feeding strategy. Could be used for inferring indoor/outdoor systems. Useful for comparing farm land use footprints
Production Area	~~Farm Agricultural Fraction~~Production area	ha/ ha	Agricultural Area / Farm Area	Measures proportion of farm used for agricultural production.
Tenure	~~Owner~~ Tenure ~~Fraction~~	ha/ ha	Owner Occupied Area / Agricultural Area	Measure of ownership structure and socio-economic characterisation.
Replacement Rate	~~Replacement Rate~~Heifers	qty/ qty	Heifers / Dairy Cows	Measure of non-productive herd

RESULTS

Relationships among KPIs

The extracted time series from the FBS were used to compute KPIs that describe dairy farms in a 16-dimensional system (see Figure S1 for distribution). Annual PCAs were computed as well as a calculation that includes all data from 2001 to 2014 (Fig ~~1~~2). Three dimensions of the PCA (PC1, PC2, and PC3) including all data sets from 2001 to 2014 ~~(Fig 1)~~ explain approximately 50% of variation (Figure S2). The loadings on the first 3 components broadly represent seven groups of KPIs (correlated in at least two components): i) milk production specifically (dairy cows, milk yield, concentrate feed per LU, and milk premium), ii) intensity and specialisation of livestock production (dairy stocking density, livestock density, dairy fraction, labour, and fodder per LU), iii) grazing prevalence (cash crop and grass presence), iv) grass/forage maize mix, v) production area, vi) tenure, and vii) replacement rate.

Fig ~~1~~2 – PCA results for all ~~KPI~~Key Performance Indicator values across all years (2001-2014). Panels on the left show the PCA scores for individual farms, on the right loading for individual metrics.

Area of land tenured by the owner of a farm is inversely related to dairy production area and replacement rate, which indicates that more heterogeneous farms with low replacement rates are more likely than more specialised dairy farms to be tenured by their owners (Fig ~~1~~2).

The component scores in Fig. ~~1~~2 (left-hand plots) show that the majority of farms are concentrated at the centre of the axes for all years (2001-2014) with some outliers for years before 2006. There is some indication that there is less diversity in farms (points are closer together) in later years.

Procrustes rotation of the first 3 components (Fig ~~2~~3) illustrates that in the periods 2001-2004 and 2006-2014 there are no large differences in the configuration of annual KPIs (sum of squares close to zero)

while 2005 appears an outlier from all other years. This result suggests that the relationship between KPIs has largely remained stable over time, suggesting that they are reliable measures of farm properties even when structural changes are occurring.

Fig ~~2~~3 - Procrustes analysis of annual variation in relationships among Key performance Indicators (KPIs) are derived from principle component analysis of annual data over the years 2001 – 2014, based on the sum of squared distances.

Cluster identification

Clustering analysis results ~~indicates~~indicate the number, configuration, and distinctiveness (mixing probabilities) of clusters for each of the survey years. Different cluster configurations are represented by the model i.e~~.,~~. VVV ellipsoidal, varying volume, shape, and orientation and VEV: ellipsoidal, equal shape. Number of farms decreased in the 14 years of study with the majority of farms distributed in mainly two or three clusters (higher probability). Further, clustering analysis identified three clusters for most years except for 2001 and 2003, which had four clusters, and 2011, 2012, and 2014, ~~that~~which had two clusters (Table 2). The distribution of farms among clusters was fairly even in most years with the exception of the smaller clusters (mixing probability < 0.1) (Table 2 & Fig 3). It is likely that these fluctuations in the smaller clusters are a combination of: (i) sampling artefacts where relatively rare farm configurations drop in and out of the sample~~,~~; (ii) farms that are in transition, or: (iii) farms that have been affected by extreme events. In the majority of years, the individual clusters varied in volume, shape and orientation (VVV) although in a few years (2007, 2009, & 2010) clusters had equal shape (VEV) (Table 2).

Table 2 – Clustering analysis results, indicating the number, configuration and distinctiveness (mixing probabilities) of clusters for each of the survey years.

Year	Cluster configuration	Number of clusters	log likelihood	n	df	Mixing probabilities			
						1	2	3	4
2001	VVV	4	1611	724	611	0.22	0.23	0.35	0.20
2002	VVV	3	431	678	458	0.50	0.48	0.02	
2003	VVV	4	862	643	611	0.38	0.30	0.30	0.02
2004	VVV	3	-182	512	428	0.48	0.37	0.16	
2005	VVV	3	32	477	458	0.42	0.52	0.06	
2006	VVV	3	393	464	458	0.42	0.35	0.23	
2007	VEV	3	67	469	428	0.46	0.42	0.12	
2008	VVV	3	337	493	458	0.55	0.42	0.03	
2009	VEV	3	366	488	428	0.47	0.44	0.09	
2010	VEV	3	623	479	428	0.40	0.15	0.45	
2011	VVV	2	390	479	305	0.37	0.63		
2012	VVV	2	454	467	305	0.44	0.56		
2013	VVV	3	1122	455	458	0.48	0.39	0.12	
2014	VVV	2	505	432	305	0.56	0.44		

Because the clustering analysis was performed for each year separately (a total of 41 clusters in 14 years of study, Table 2), we computed a MDS across clusters from all years in order to track the same type of cohort from one year to another. MDS in a single dimension had an almost perfect linear fit (R=0.999) between the ordination distance and the observed dissimilarity. Thus, this second classification allows changes to be tracked over time for the same type of cluster, and the comparison of clusters. The results show two predominant clusters through time (groups 1 & 2) and other smaller groups (Fig ~~3~~4). Since there are two main clusters from 2001 to 2014 and four more clusters that appear infrequently, we focused the following discussion on the two predominant clusters: group 1 (circles), classified as "*extensive systems";* group 2 (triangles), categorized as "*intensive systems"* (Fig ~~3~~4). Note that the outlying groups could comprise very intensive or "mega dairy" farms (Fig ~~3~~4). All clusters are coloured by their MDS value; so similar clusters can be tracked over time and separated from less common typologies (Fig ~~3~~4).

Fig ~~3 – KPIs trends~~4 – Trends in mean ~~KPI~~Key Performance Indicator values for all identified clusters over the period 2001 – 2014~~, with the~~. The number of farms in each cluster ~~are~~is represented by the size of symbol. Intensive systems are represented by triangles and extensive systems by circles. The solid black line represents the KPI annual average. The distance among all clusters in all years of study is represented by the colour scale MDS. This distance allows identifying which clusters are more similar.

Cluster comparison and trends

KPIs such as number of dairy cows, milk yield, ~~concentrates~~concentrate use, grass in agricultural area, as well as productive area have increased in the 14 years of study (Fig ~~3~~4); while the other KPIs have remained fairly constant, except for labour intensity, where lower labour use per hectare has been observed over time. Note that the black ~~steady~~ line that represents the KPI annual average at sectoral level rarely explains the actual value of a cluster for a particular year; extensive and intensive systems are generally above or below.

Fig ~~3b~~4b shows the difference in milk production between clusters. The intensive farms consistently produced more milk per cow than extensive systems. In 2014, there was an annual difference of 665 l/dairy cow in productivity and 8 hours/ha in labour intensity (Fig ~~3j~~4j). Intensive systems had an additional 40 dairy cows (Fig 3a) that consume 224 kg/yr more concentrate per LU (Fig ~~3c~~4c), 21 kg/yr less coarse fodder per LU (Fig ~~3i~~4i), and had a higher ratio of fodder and maize to grass (Fig ~~3n~~4n-o) (Table 3). There was no consistent difference in milk price premiums between intensive and extensive systems over the study period (Fig ~~3d~~4d), though intensive systems fared slightly better for more years.

Despite the difference in milk production, the replacement rate (Table 3), inferred from heifer to cow ratio (Fig 3e), has remained similar for intensive and extensive farms over the 14 years under study, initially declining for both clusters before increasing again towards a peak in 2012. Intensive farms have higher overall stocking densities (Fig ~~3g~~4g-h), but not necessarily a higher dairy fraction (3f) than extensive farms, though differences are small. In 2014, both systems had 1.4 dairy cows per hectare and 2 LU/ha (Table 3). Intensive and extensive systems can also be differentiated by the utilization rate of non-cash crop area (Fig ~~3l~~4l) and grass (Fig ~~3m~~4m) in the agricultural area of a farm, which have not changed dramatically since 2001 for each system (Table 3). Extensive systems utilised almost all non-cash crop agricultural area for grass production, compared to more intensive systems that only used 70% for grass production (Fig ~~3l~~4l-m and Table 3). Intensive farms produced maize on an area equivalent to 20% of grass area, and included fodder areas that grew over time, whilst extensive farms did not produce maize and are characterised by very small fodder areas compared to their grass extent (Fig ~~3o~~4o). There are a small number of intensive farms represented in the sporadic small clusters that appear in some years with comparatively very large maize areas (Fig ~~3o~~4o). Between 2001 and 2014, productivity (l cow^{-1}) in extensive systems increased by 17% and cow numbers by 52%, compared with increases of 13% and 17% in productivity per cow and cow numbers, respectively, for intensive systems (Table 3). At the same time, labour intensity (hours/ha) has declined by 17-18% and the use of

concentrate feed has increased by 500-600 kg/LU/yr in both intensive and extensive systems over the study period (Table 3). Across all farm types there was a large increase in the use of concentrates between 2005 and 2008 (Fig 3c).

Extensive systems were characterised by a 10% lower rate of owner occupation in 2001, which converged to a similar rate as for intensive systems in 2014, at around 60% of tenure (Fig 4k) (Table 3). Non-agricultural area, such as woodland, buildings, roads, water, and household gardens account for just 3% of farm areas across both intensive and extensive systems (Table 3).

Table 3 – Comparison of Key Performance Indicators (KPIs) between 2001 and 2014 for extensive (E) and intensive (I) farm cluster.

KPIs	E 2001	E 2014	I 2001	I 2014	E[2014 -2001]	I[2014-2001]	2001 [E-I]	2014 [E-I]
~~Dairy Cows (qty)~~Total dairy cows	78	132	145	172	55	27	68	40
Milk ~~Yield (l/dairy cow)~~yield	5,784	6,835	6,588	7,499	1,051	911	-804	- 665
Milk ~~Price Premium (£/l, £/l)~~premium	1.01	0.98	1.04	0.97	-0.03	-0.07	-0.03	-0.01
~~Concentrated Feed (tonne/LU)~~Concentrate fed	0.77	1.29	0.91	1.52	0.52	0.61	-0.13	- 0.22
~~Coarse Fooder (tonne/LU)~~Fodder fed	0.25	0.16	0.25	0.14	-0.09	-0.11	0.00	-0.02
~~Dairy Cows Fraction (qty/LU)~~Cow fraction	0.6	0.6	0.6	0.6	0.1	0.0	0.0	0.0
~~Livestock Density (LU/ha)~~Cow stocking rate	1.9	2.0	2.1	2.1	0.1	0.0	-0.2	-0.1
~~Dairy Stocking Density (dairy cows/ha)~~Livestock density	1.2	1.4	1.4	1.4	0.1	0.0	-0.2	0.0
Labour ~~Intensity (hours/ha)~~intensity	82	68	71	60	-14	-11	10	8
Fodder ~~Grass Ratio (ha/ha)~~area	0.0	0.0	0.2	0.3	0.0	0.1	-0.2	-0.3
~~Maize~~ Grass ~~Ratio (ha/ha)~~area	0.0	0.0	0.2	0.2	0.0	0.1	-0.2	-0.2
~~Fraction of~~ Non-~~Cash Crop Area~~cash crop area in ~~Agricultural Area (ha/ha)~~agricultural area	1.0	1.0	0.9	0.8	0.0	0.0	0.1	0.1
~~Grazing Agricultural Area Ratio (ha/ha)~~Grass area in agricultural area	1.0	1.0	0.7	0.7	0.0	-0.1	0.3	0.3
~~Farm Agricultural Fraction (ha/ha)~~Production area	0.97	0.97	0.98	0.97	0.00	0.00	0.00	0.00
~~Fraction Owner Occupied Area (ha/ha)~~Tenure	0.6	0.6	0.5	0.6	0.0	0.0	0.1	0.0
~~Replacement Rate (heifers/cows)~~Heifers	0.2	0.2	0.2	0.2	0.0	0.0	0.0	0.0

Formatted: Font color: Text 1

Formatted: Font color: Text 1

Formatted: Font color: Text 1, English (United Kingdom)

Formatted: Font color: Text 1, English (United Kingdom)

DISCUSSION

Identifying farm typologies and trends

Analysis of all dairy farm FBS data for the years 2001 through to 2014 confirms a trend of dairy consolidation (3), with a 41% decline in the number of dairy farms surveyed over that time period, and an 18% decline in dairy cow numbers - in line with separate statistics showing that the dairy cow population in England & Wales has declined by 19% over the same period. The total area represented by FBS dairy farms declined by just 7%, reflecting an increase in average dairy farm size from 132 ha in 2001 to 141 ha in 2014. Intensification is demonstrated by the 13-17% increase in milk production per cow over the same period. While the FBS data are drawn from a broadly representative set of farms recruited from agricultural census data there will inevitably sampling issues leading to the possible omission of rarer farm types such as very large and intensive dairy farms and farms following a more extensive pathway ~~should~~if these businesses ~~decline the invitation~~choose not to ~~take part in the FBS~~be surveyed.

~~Many of the 16 KPIs evaluated changed significantly over the 14-year study period, reflecting productivity and efficiency improvements linked to diet change and technological advances. For example, labour use has generally declined across all farms possibly representing a technology advancement or employing more specialized labour. Consolidation is reflected in declining numbers of farms over a slightly declining aggregate area of dairy farms. Intensification is reflected in higher milk yields per cow, higher stocking rates and increased rates of fodder and concentrate feeding over time, coupled with a decline in labour intensity from 2001 to 2014 that presumably reflects technological improvements (e.g. investment in more efficient and automated milking parlours) and increased herd size linked to higher performing businesses remaining in dairy; this is in line with previous findings that have demonstrated the importance of efficient labour use in dairy (48,49).~~

Many of the 16 KPIs evaluated changed significantly over the 14-year study period, reflecting productivity and efficiency improvements linked to diet change and technological advances. Consolidation is reflected in declining numbers of farms over a slightly declining aggregate area of dairy farms. Intensification is reflected in higher milk yields per cow, higher stocking rates and increased rates of fodder and concentrate feeding over time, coupled with a decline in labour intensity from 2001 to 2014 that presumably reflects technological improvements (e.g. investment in more efficient and automated milking parlours) and increased herd size linked to higher performing businesses remaining in dairy; this is in line with previous findings that have demonstrated the importance of efficient labour use in dairy (50,51).

We employed a methodology to characterise dairy farm typologies in England and Wales, underpinned by transformation of economic metrics reported in the FBS into KPIs that reflect physical and socio-economic characteristics of dairy farms. While similar approaches have been used previously our methodology is particularly robust because it combines multivariate analyses (PCA) with a Gaussian mixture model-based clustering and multidimensional scaling for grouping similar clusters over time. The method provides robust guidance on the number of clusters to choose. We found that the more usual clustering approaches (k-mean and k-medoid) did not perform well on these data which suggests their use may be inappropriate on FBS and similar data, especially when clusters are not expected to be of equal size or shape. This method can be applied to all agricultural sectors in all countries where farm economic statistics are compiled, providing a solid basis for rigorous and comprehensive evaluation of the sustainability of farm typologies and identified intensification pathways. This approach provides a robust basis for modelling the sustainability of pathways of intensification through time. For example, attributional LCA can be applied to determine the environmental footprints of milk production for statistically defined dairy farm typologies within (e.g. ~~(50)~~(52)) and between (e.g. ~~(51)) years.~~(53)) years. Consequential LCA may also be applied to evaluate the environmental loading changes that arise,

directly and indirectly, when farm typologies evolve in characteristics and predominance over time ~~(52).~~(54).

Clustering analysis of KPIs identified two main dairy farm typologies, representing different levels of intensification and following different but somewhat convergent pathways of intensification between 2001 and 2014. Results indicate that the dairy sector was more heterogeneous in the earlier years of the study, comprising three or four distinct clusters of dairy farms, and consolidated into just two main types of farm that predominated since 2011 (except for 2013). Classification of the two predominant farm typologies as "extensive" and "intensive" may appear to be a simplification, but these types are statistically identified and this does concur with a previous industry report ~~(53)~~(55) that identified two profitable pathways of dairy farm business development: (i) grass-based expansion to maximise margins per litre of milk; (ii) farm intensification to maximise margins per hectare. Intensive farms achieve higher milk yields per cow~~,~~ but use more concentrate feed and maize than extensive farms, which rely more heavily on grass and do not use maize. Smaller groups of farms identified in some years by clustering analysis had some characteristics indicative of very intensive farming methods, and may reflect a subset of emerging "mega-dairy" farms within the intensive cluster. Results also indicate some degree of convergence between the two main farm typologies, owing to a faster pace of intensification (e.g. increasing milk yield per cow) among the extensive farms. Consequently, in 2014, intensive and extensive farms had the same stocking densities per hectare.

Evaluating the sustainability of intensification

Ongoing C&I is shifting UK dairy production away from small and medium sized farms towards larger farms that can be broadly categorised as grazing- or indoor- dominated systems. C&I pathways influence animal diets, health, yields, grassland and manure management, with implications for environmental and economic efficiency at animal, farm and system level. Whilst the definition of sustainable

intensification is contested and may have different meaning in different contexts ~~(54)~~(56), a broad definition is to raise productivity and social welfare while reducing environmental impacts per unit of output. The measures captured through the use of farm survey data only ~~capture~~include a small subset of those in a recent meta-analysis (14). A more complete analysis would require socio-economic, biodiversity and soil health indicators. There is some evidence that environmental and economic indicators may be correlated but social indicators differ ~~(55,48).~~(50,57). However, any regional or national analysis upscaling from farms requires being able to identify typologies of ~~farm types~~farms for which these indicators could be collected~~.~~ in a targeted manner. The indicators developed here ~~that~~ can be linked to environmental performance, for example, feed strategies and land use that embody upstream land and environmental impact (e.g~~.,~~. land use, resource depletion, GHG emissions). Therefore, results of this research can be used to model scenarios including social aspects (e.g~~.,~~. labour intensity), economic components (e.g. profits per litres of milk), and environmental impacts (e.g~~.,~~. carbon, land and water footprints) of dairy farming. The clusters also provide a more accurate profile of trends in the sector than hitherto provided by analysis of "average" farms or aggregate data. There is also potential for application in terms of farm management as the developed KPIs could also be used to benchmark farms within cluster typologies, for example in terms of feed use efficiency, and to recommend priority practises to sustainably intensify that are targeted to the distinct cluster typologies.

We show that the UK dairy sector can be characterised by 2-3 clusters over the period of study, which allows environmental footprints to be readily calculated using LCA methods. Notably, across all clusters concentrate feed use, and by implication the indirect land footprint of dairy production, increased. A comprehensive analysis would require a wider system boundary than the individual farm. Increased maize and concentrate feed has the potential to improve animal-level efficiency and reduce on-farm environmental footprints ~~(56,57), but may not reduce system level footprints owing to possible land use change GHG emissions (56–58). Coupling the evolution of farm typologies described here with feed~~

~~sourcing statistics and dairy-beef production models would enable a full LCA appraisal of direct and indirect environmental consequences arising from changes to animal diets and beef co-production (52).~~(58,59), but may not reduce system level footprints owing to possible land use change GHG emissions (58–60). Coupling the evolution of farm typologies described here with feed sourcing statistics and dairy-beef production models would enable a full LCA appraisal of direct and indirect environmental consequences arising from changes to animal diets and beef co-production (54). Findings from this study may be directly transferable to dairy farming in other industrialised countries where similar C&I trends prevail.

~~Thus, this research provides a foundation for further analyses, in particular LCA (50,59) and DEA (60,61) that can address all aspects of sustainable intensification. Such work could represent a significant advance on previous studies of dairy intensification that have primarily focussed on environmental (50,51) or socio-economic factors~~ Thus, this research provides a foundation for further analyses, in particular LCA (52,61) and DEA (62,63) that can address all aspects of sustainable intensification. Such work could represent a significant advance on previous studies of dairy intensification that have primarily focussed on environmental (52,53) or socio-economic factors (9,17–20) in isolation.

CONCLUSION

Trends in dairy farm intensification are usually reported at the sectoral or "average farm" level, sometimes differentiated into regions or percentage quartiles, in terms of economic outputs, inputs, and margins ~~(62)~~(64). Although useful for detecting broad trends, this approach does not adequately capture heterogeneity in farm operations, and does not relate business structure to the physical characteristics necessary to fully evaluate the sustainability of intensification trends. ~~We show that~~In the method developed here indicators calculated from farm business survey data coupled with robust model-based clustering identify the number of groups of farms and trends over time. We show that in

England and Wales dairy farms have largely consolidated and specialized into two distinct clusters that now predominate within the sector: one “extensive” cluster of farms relying on expansion of grass-based milk production, with lower milk yields and labour intensity; one “intensive” cluster of farms producing, on average, more milk per cow with more concentrate and more maize, but fewer hours of labour per hectare. There is some indication that these clusters are converging as the extensive cluster is intensifying slightly faster than the intensive cluster, in terms of milk yield per cow and use of concentrate feed. ~~These~~The statistical characterisation of these groups will allow the accurate evaluation of the consequences of dairy C&I at national and international scales to be advanced.

ACKNOWLEDGEMENTS

The authors acknowledge funding provided by the Sêr Cymru National Research Network for Low Carbon Energy and Environment (NRN-LCEE~~).~~) and the comments of three anonymous referees.

REFERENCES

1. Steinfeld H, Gerber P, Wassenaar T, Castel V, Rosales M, De Haan C. Livestock’s Long Shadow: Environmental Issues and Options. FAO, Rome, Italy,. 2006.

2. Steinfeld H, Gerber P. Livestock production and the global environment: Consume less or produce better? Proc Natl Acad Sci. 2010;107(43):18237–8.

3. March MD, Haskell MJ, Chagunda MGG, Langford FM, Roberts DJ. Current trends in British dairy management regimens. J Dairy Sci. 2014;97(12):7985–94.

4. AHDB Dairy. Dairy statistics: an insiders guide 2016. 2016.

5. AHDB Dairy. Producer numbers for England and Wales [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/market-information/farming-data/producer-numbers/uk-producer-

numbers/

6. AHDB Dairy. UK cow numbers [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/market-information/farming-data/cow-numbers/uk-cow-numbers/

7. AHDB Dairy. Average UK Milk Yield [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/market-information/farming-data/milk-yield/average-milk-yield/#.WjpZvxOFiOE

8. European Commission. Structure and dynamics of EU farms : changes , trends and policy relevance. EU Agricultural Economics Briefs. 2013;(9):1–15.

9. Alvarez A, del Corral J, Solís D, Pérez JA. Does Intensification Improve the Economic Efficiency of Dairy Farms? J Dairy Sci. 2008;91(9):3693–8.

10. Van Grinsven HJM, Ten Berge HFM, Dalgaard T, Fraters B, Durand P, Hart A, et al. Management, regulation and environmental impacts of nitrogen fertilization in northwestern Europe under the Nitrates Directive; A benchmark study. Biogeosciences. 2012;9(12):5143–60.

11. Dairy Development Centre. Improving the Welsh dairy supply chain: Grass value. Camarthen; 2014.

12. Veysset P, Lherm M, Roulenc M, Troquier C, Bébin D. Productivity and technical efficiency of suckler beef production systems: Trends for the period 1990 to 2012. Animal. 2015;9(12):2050–9.

13. Bava L, Sandrucci A, Zucali M, Guerci M, Tamburini A. How can farming intensification affect the environmental impact of milk production? J Dairy Sci. 2014;97(7):4579–93.

14. Mahon N, Crute I, Simmons E, Islam MM. Sustainable intensification – 'oxymoron' or 'third-way'? A systematic review. Ecol Indic. 2017;74:73–97.

15. Garnett T, Appleby MC, Balmford A, Bateman IJ, Benton TG, Bloomer P, et al. Sustainable Intensification in Agriculture: Premises and Policies. Science (80-). 2013;341(6141):33–4.

16. Garnett T, Godfray C. Sustainable intensification in agriculture. Navigating a course through competing food system priorities. Food Climate Research Network and the Oxford Martin Programme on the Future of Food. 2012.

17. Cabrera VE, Solís D, del Corral J. Determinants of technical efficiency among dairy farms in Wisconsin. J Dairy Sci. 2010;93(1):387–93.

18. Dong F, Hennessy DA, Jensen HH, Volpe RJ. Technical efficiency, herd size, and exit intentions in U.S. dairy farms. Agric Econ. 2016;47(5):533–45.

19. Toma L, March M, Stott AW, Roberts DJ. Environmental efficiency of alternative dairy systems: A productive efficiency approach. J Dairy Sci. 2013;96(11):7014–31.

20. Brotzman RL, Cook NB, Nordlund K, Bennett TB, Gomez Rivas A, Döpfer D. Cluster analysis of Dairy Herd Improvement data to discover trends in performance characteristics in large Upper Midwest dairy herds. J Dairy Sci. 2015;98(5):3059–70.

21. Sturaro E, Marchiori E, Cocca G, Penasa M, Ramanzin M, Bittante G. Dairy systems in mountainous areas: Farm animal biodiversity, milk production and destination, and land use. Livest Sci. 2013;158(1–3):157–68.

22. Maseda F, Díaz F, Alvarez C. Family dairy farms in Galicia (N.W. Spain): Classification by some family and farm factors relevant to quality of life. Biosyst Eng. 2004;87(4):509–21.

23. Duchy College Rural Business School. Farm Business Survey, 2011-2012: Special Licence Access. [data collection]. 2nd Edition. 2014.

24. Duchy College Rural Business School. Farm Business Survey, 2013-2014: Special Licence Access. [data collection]. 3rd Edition. 2016.

25. Duchy College Rural Business School. Farm Business Survey, 2012-2013: Special Licence Access. [data collection]. 2nd Edition. 2015.

26. Duchy College Rural Business School. Farm Business Survey, 2014-2015: Special Licence Access. [data collection]. 2nd Edition. 2016.

27. Defra, National Assembly for Wales. Farm Business Survey, 2001-2002: Special Licence Access. [data collection]. 2004.

28. Defra, National Assembly for Wales. Farm Business Survey, 2002-2003: Special Licence Access. [data collection]. 2004.

29. Defra, National Assembly for Wales. Farm Business Survey, 2003-2004: Special Licence Access. [data collection]. 2005.

30. Defra, National Assembly for Wales. Farm Business Survey, 2005-2006: Special Licence Access. [data collection]. 3rd Edition. 2008.

31. Defra, National Assembly for Wales. Farm Business Survey, 2006-2007: Special Licence Access. [data collection]. 2008.

32. Defra, National Assembly for Wales. Farm Business Survey, 2007-2008: Special Licence Access. [data collection]. 2010.

33. Defra, National Assembly for Wales. Farm Business Survey, 2008-2009: Special Licence Access. [data collection]. 2nd Edition. 2011.

34. Defra, National Assembly for Wales. Farm Business Survey, 2009-2010: Special Licence Access.

[data collection]. 3rd Edition. 2014.

35. Defra, National Assembly for Wales. Farm Business Survey, 2010-2011: Special Licence Access. [data collection]. 3rd Edition. 2014.

36. Defra, National Assembly for Wales. Farm Business Survey, 2004-2005: Special Licence Access. [data collection]. 2006.

37. AHDB Dairy. Feed prices [Internet]. 2017. Available from: https://dairy.ahdb.org.uk/resources-library/market-information/farm-expenses/feed-prices/#.WBi26-GLRR4

38. Defra. Calendar year farmgate milk prices UK 1970 to 2014 [Internet]. 2016. Available from: https://data.gov.uk/dataset/milk_prices_and_composition_of_milk/resource/42f5e421-b6db-45a4-a5e0-08f65c040413

39. Farm business survey - technical notes and guidance [Internet]. Available from: https://www.gov.uk/guidance/farm-business-survey-technical-notes-and-guidance

40. AHDB Dairy. Evidence Report: GB dairy herd performance 2014/15. Kenilworth; 2015.

41. Redman G. The John Nix Farm Management Pocketbook 2012. 40th ed. Melton Mowbray: The Andersons Centre; 2011. 280 p.

42. SAC Consulting. The Farm Management Handbook. 30th ed. McBain C, Curry J, editors. Edinburgh: SAC Consulting; 2016. 332 p.

43. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. Data Mining, Inference, and Prediction. 2nd ed. Springer-Verlag; 2009. 763 p.

44. Mardia KV, Kent JT, Bibby JM. Multivariate Analysis. Academic Press; 1979.

45. R Core Team. R: A language and environment for statistical computing. Vienna, Austria; 2017.

~~44~~46. Oksanen J, F. Guillaume, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package. R package version 2.0-2. 2011.

~~45~~47. Fraley C, Raftery AE, Murphy TB, Scrucca L. mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Technical Report 597, University of Washington. 2012.

~~46~~48. Fraley C, Raftery AE. Model-based clustering, discriminant analysis, and density estimation. J Am Stat Assoc. 2002;97(458):611–31.

~~47~~49. Fraley C, Raftery AE. Bayesian Regularization for Normal Mixture Estimation and Model-Based Clustering. J Classif. 2007;24(2):155–81.

~~48~~50. Soteriades AD, Faverdin P, March M, Stott AW. Improving efficiency assessments using additive data envelopment analysis models: an application to contrasting dairy farming systems. Agricultural and Food Science. 2015;24(3):235–48.

~~49~~51. Wilson P. Decomposing variation in dairy profitability: The impact of output, inputs, prices, labour and management. J Agric Sci. 2011;149(4):507–17.

~~50~~52. O'Brien D, Shalloo L, Patton J, Buckley F, Grainger C, Wallace M. A life cycle assessment of seasonal grass-based and confinement dairy farms. Agric Syst. 2012;107:33–46.

~~51~~53. Capper JL, Cady RA, Bauman DE. The environmental impact of dairy production: 1944 compared with 2007. J Anim Sci. 2009;87(6):2160–7.

~~52~~54. Styles D, Gonzalez-Mejia A, Moorby J, Foskolos A, Gibbons J. Climate mitigation by dairy intensification depends on intensive use of spared grassland. Global Change Biology. 2017;

~~53~~55. The Andersons Centre. The structure of the GB dairy farming industry – what drives change?

2013.

~~54~~56. Loos J, Abson DJ, Chappell MJ, Hanspach J, Mikulcak F, Tichit M, et al. Putting meaning back into 'sustainable intensification'. Front Ecol Environ. 2014;12(6):356–61.

~~55~~57. Reig-Martínez E, Gómez-Limón JA, Picazo-Tadeo AJ. Ranking farms with a composite indicator of sustainability. Agric Econ. 2011;42(5):561–75.

~~56~~58. Vellinga T V., Hoving IE. Maize silage for dairy cows: Mitigation of methane emissions can be offset by land use change. Nutr Cycl Agroecosystems. 2011;89(3):413–26.

~~57~~59. Van Middelaar CE, Berentsen PBM, Dijkstra J, De Boer IJM. Evaluation of a feeding strategy to reduce greenhouse gas emissions from dairy farming: The level of analysis matters. Agric Syst. 2013;121:9–22.

~~58~~60. Styles D, Gibbons J, Williams AP, Stichnothe H, Chadwick DR, Healey JR. Cattle feed or bioenergy? Consequential life cycle assessment of biogas feedstock options on dairy farms. GCB Bioenergy. 2015;7(5):1034–49.

~~59~~61. Styles D, Gibbons J, Williams AP, Dauber J, Stichnothe H, Urban B, et al. Consequential life cycle assessment of biogas, biofuel and biomass energy options within an arable crop rotation. GCB Bioenergy. 2015;7(6):1305–20.

~~60~~62. Iribarren D, Hospido A, Moreira MT, Feijoo G. Benchmarking environmental and operational parameters through eco-efficiency criteria for dairy farms. Sci Total Environ. 2011;409(10):1786–98.

~~61~~63. Wilson P, Lewis M, Crane R, Robertson P, Bonner J, Davenport R, et al. Farm Level Performance : Identifying Common Factors Determining Levels of Performance. 2012.

~~62~~64. Mchoul H, Robertson P, Wilson P. Dairy Farming in England. 2011.

As in the previous submission we thank both reviewers for their helpful efforts and input. For clarity in the point-by-point response we have deleted any comments and responses that relate to the previous revision and retained only those where issues remained. Our reply is in red with quotes from the revised manuscript in italic.

Reviewer #1's suggestions were all suggested edits and we have made all the changes suggested and reread through the manuscript for English as suggested

Reviewer #2's suggestions while recognising that the manuscript has been improved asks for further statistical explanation as well as suggested edits. We have done all the suggested edits and further explained the statistical approach. The main addition is Figure 1 which shows the workflow of the statistical analysis. This was a great suggestion and really improves the manuscript. We have also added further written explanation to the Data Analysis section. Reviewer #2 also asked that we illustrate the methods with a simplified data set, for space and emphasis reasons we have not done this within the manuscript but point to a book with a freely available PDF version:
https://web.stanford.edu/~hastie/ElemStatLearn/
that illustrates the approaches with simple data sets.

Reviewer #1
General comments
The manuscript has improved but it needs some other changes. In my opinion, it can be published after minor revision.
Firstly I consider very important to improve the aim and the conclusion according to the new title
Moreover, after the first revision, some new or revised sentences rather unclear and several typos appeared. I suggest rereading accurately the entire manuscript to amend the errors and clarify the meaning of some phrases.
Finally some comments were only partially addressed

Specific comments

Title
Thank you for having modified the title. I can suggest further optional alternatives:
1) Metrics for characterizing Dairy Farm Intensification: the case of England & Wales Between 2001 And 2014
2) Metrics for characterizing Dairy Farm Intensification
This is because the methods proposed is also applicable to other time periods

Thank you for the suggestion you make a good point about time so we have changed to *"Metrics and methods for characterizing dairy farm intensification using farm survey data"*

Abstract
Line 10: Please, replace the 2nd trend with a synonym in the first sentence: Evaluation of

the sustainability of trends in agricultural intensification requires comprehensive analysis of trends in farm performance across physical and socio-economic aspects

Changed the first to "changes"

Introduction
Line 35: 5% of global anthropogenic gas emissions?
Done

Lines 38-39: The sentence is not so clear.
Changed to *"One route to this is to reduce land-use intensity of milk production by increasing milk yields per cow (2)."*

Lines 39-41: I do not understand the meaning of the sentence

Changed to *"However, without advances in technology an environmental gain will only be achieved if the increase in production per cow out paces the increase in demand."*

Lines 44-46: I suggest reversing the sentences: The UK dairy industry is 10th largest global producer of cow milk (accounting for 2.2% of world production) (5) and is an exemplar of worldwide intensification trends.
Done

Line 59: I suggest avoiding double round brackets. You can use a comma or a semicolon before the number of the citation
I would prefer to keep as it consistently differentiates a literal number from a citation, however we are happy to have this formatted as per editor or journal style preferences if this is required.

Line 61: because despite. Are they both useful?
Inserted a comma between "because, despite"

Lines 117-127: please rewrite and refine the aim of the study
This is methods rather than the aim but we have removed "ratios, densities and intensities"

Methods
Lines 91-93: I suggest conjugating verbs to the past simple (identified…investigated)
Done

Line 125: to provide
Done

Results
I still see a lack of consistency among the names of indicators in table 1, fig. 1, fig 3, In the text and now also in fig s1 in the supplementary material.
Thanks, made much more consistent, though still retained some extra words in Table 1 to avoid ambiguity when the units are not presented

Line 185: I suggest deleting the reference to fig 1 because it has just been recalled in the previous line
Done

Fig. 1
In the figure title please put "Key performance indicators" in extenso instead of KPI
Done

Despite the improvements in the readability of indicators names in the PCAs, I can still count only 15 indicators in the first two PCAs and 16 in the third. The indicator Non-cash crop area ha ha-1 seems to appear only in the last PCA
Thanks for the reminder about this, the scale on the axis was omitting the final variable, now adjusted so the scale is -0.45 to 0.45 on all three graphs.

Table 3
Replacement rate seems in bold and/or in a different font.
Well spotted, 10 point rather than 9 point, fixed

Fig 3
Please, check the English of the figure title
In the figure title I suggest putting "Key performance indicators (KPI)" in extenso instead of KPI, the first time
Changed to "***Fig 4 –*** *Trends in mean Key Performance Indicator values for all identified clusters over the period 2001 – 2014. The number of farms in each cluster is represented by the size of symbol. Intensive systems are represented by triangles and extensive systems by circles. The solid black line represents the KPI annual average. The distance among all clusters in all years of study is represented by the colour scale MDS. This distance allows identifying which clusters are more similar."*

Discussion
Lines 293-296: Check the English of the phrase and, if possible, break into shorter sentences.
Changed end to *"if these businesses choose not to be surveyed."*

Lines 345: avoid if possible the repetition of capture
Changed the second to "include"

Lines 349-352: Check the English of the phrase
Removed "that"

Conclusions
I think conclusion section needs an improvement. The conclusion focuses only on the specific results of the England and Wales case. Conclusion about the method and its contribution to research are missing.
Added *"In the method developed here indicators from farm business survey data coupled with robust clustering identify groups of farms and trends over time."*

Line 381: I suggest adding the specification “in England and Wales”
Done

Reviewer #2: I would like to thank the authors for their efforts. This new version of the manuscript has been significantly improved. However, I am afraid that the manuscript is still not ready for publication, because the methodology remains insufficiently explained. Answers to many of my comments are incomplete or have been completely ignored. I have made great efforts in this and the previous version to help the authors improve the presentation of their methods and would like to see my recommendations materialize. I think that this study is very interesting so I am only asking the authors to take my recommendations into account to make the manuscript more accessible to readers.

Please find my comments below. Every comment starts with ‘RESPONSE:’

l.66: Sentence starting with ‘While’ does not make sense. I think you meant to link this sentence with the previous one as follows: ‘[…] (8), while Alvarez et al. (9) […]’. This would result to a very long sentence. I recommend that you just start the second sentence as ‘Alvarez et al. (9) […]’.
Done

l:59:82: RESPONSE: Thank you for the edits. Choice of French beef example fair enough. Only comment here is that I would like to see a better linking between the following sentences: ‘It might be expected that intensification of dairy farms will result in more efficient farms growing at the expense of less efficient farms. However, evidence from the livestock sector suggests that farms may become less economically efficient because, despite an increase in investment in capital, technology and concentrate feed, output may remain constant over several decades, as is the case for French beef farms (Veysset et al., 2015)’.
Inserted *“Given this range in efficiency it…”*at the start of the 2nd sentence

l.122-123: A comma following ‘e.g.’ or ‘i.e.’ is unnecessary. This typo appears in several places in the manuscript. Please replace ‘e.g.,’ and ‘i.e.,’ with ‘e.g.’ and ‘i.e.’ respectively.
Thanks, eliminated all rogue commas

Statistics. RESPONSE: Thank you for your efforts. However, as with the earlier version, I think that you have not placed enough effort on satisfactorily shaping this subsection and on addressing my comments. For example, you are assuming that everyone is familiar with PCA loadings (l.215) and have not explained what they are. In fact, you have not even explained what PCA is, how it works and why you have used it. You obviously want to represent KPIs with uncorrelated variables. Explain why.
Please briefly describe PCA, it is very easy to do so. See section 4 in Jollands et al. (2004) - an excellent brief description of PCA. Please add textbook references for PCA and Procrustes rotation as no one is going to learn/understand these methods just by reading the R

documentation.
Also, it is unclear why Procrustes rotation of the loadings is necessary. That 'the sign of component loadings is arbitrary' (l.219-220) does not help the reader understand why Procrustes is used, especially given that many readers may not know what loadings are, and that they have a sign.
With k-means, clusters are not necessarily of 'equal size' (l.231)? See Alvarez et al. (2008).
My requests for a diagram illustrating how the different models/methods are combined; and for a short example demonstrating your novel approach and its advantages 'in action' seem to have been completely ignored. Right now, this section feels 'overloaded' in terms of methodology: PCA, Procrustes rotation, Gaussian mixture model-based clustering, Expectation Maximization algorithm, Maximum Likelihood, BIC, multidimensional scaling, shape, volume and orientation of multidimensional datasets… this is too much for the 'intelligent lay reader'. Personally, I am not really following. A visual summary of how all methods combined step-by-step, as well as a trivial example (perhaps visual too), are absolutely necessary.

Sorry that we did not provide more explanation in the last revision. Thank you for the suggestions here, we now include a new Figure (Figure 1) which illustrates the flow through the methods and how they relate. We have further edited the introduction to the Data Analysis section to describe PCA. Thank you for the reference suggestion, we have now inserted a reference to Hastie et al. which is a very good book, a PDF is freely available from the author (https://web.stanford.edu/~hastie/ElemStatLearn/) and it covers the majority the methods; we use the additional reference to the Mardia et al. book that covers the rest. While illustrating the methods with simple examples is a good suggestion we think that this would change the emphasis of the paper and there would not be space to do this justice. As the Hastie et al. book illustrates the methods with simple and consistent examples we don't think it is necessary to duplicate this within the manuscript.
We have expanded the explanation of Procrustes rotation to explain further this aspect. Technically you are correct to state that k-mean and related methods don't explicitly assume clusters are the same size. However, in the case where they are different sized very surprising cluster results are produced. I.e. practically the assumption holds. There are a couple of useful blog entries on this with illustrations
https://blog.learningtree.com/assumptions-ruin-k-means-clusters/
http://varianceexplained.org/r/kmeans-free-lunch/
Taking on board the recommendations, the main text edits are at the start of the Data Analysis section which now reads:
"We use a suite of statistical methods and workflow to analyse the data as shown in Fig 1. Further details of all the analysis methods with illustrations on simple data sets are available in (43,44) in particular we recommend chapter 14 of Hastie et al. All code to reproduce the data analysis is available on request from the authors. PCA (principal components analysis) was used to explore the relationship among KPIs (i.e. identification of fundamental farm properties) and how these relationships change over time. The usual aim is to reduce multiple dimensions down to two or three for illustration and analysis purposes. PCA creates new linear combinations of existing variables (components) ranked to explain as much variation as possible. The relative weighting of each KPI on each component is then termed the loading and value each farm on the component the score. For the set of KPIs to be a useful measure of farms over time, the relationship between KPIs should be relatively

constant but change should result in farms changing their position along the KPI dimensions. PCA was calculated in R (45) and Procrustes rotation of the first 3 KPI loadings identified by PCA was used to compare the structure of each year and compare structure between years with the vegan package (46). The Procrustes analysis rotated the PCA loadings to minimize the sum of squares of the difference in distance between loading for each year pair, a small total sum of squares indicating the relationship between the individual KPIs between years was similar, a large difference that the relationship changed between years. The rotation is necessary to fairly compare between years as the relationship between the variables and hence the relative loadings may remain constant over time but the absolute loadings may change and the sign of component loadings is arbitrary (can be positive or negative depending on the algorithm or data used)."

l.189: '[…] and cash crops provide information […]'. Do you want to say '[…] and cash crops that provide information […]'?
Inserted "to" to correct the grammar

l.141: Can you please confirm that the assumption of a Gaussian distribution does not impact on the validity of your analysis. How do your KPIs look like? I would recommend that you provide histograms and boxplots of your 16 indicators as supplementary information. The shape of the clusters is quite flexible in the sense they can range from speherical to long and thin, though (as you would expect in a cluster) could not be discontinuous or not smooth.. As indicated in Table 2 the majority of clusters were of the 'VVV' type indicating that this flexibility was necessary.
RESPONSE: Apologies for my limited understanding, but I do not see how this answers my question about the assumption of normality.
Possibly we misunderstood you here in relation to the initial suggestion. The boxplots in the supplementary material included in the last revision did illustrate the distribution of the variables. Non-Gaussian distributions would be accommodated because e.g. in the case of a multimodal distribution, there could be a cluster at each mode in that dimension. I.e individual clusters are multivariate Gaussian though not constrained to be spherical, but this does not constrain the underlying variables to be Gaussian.

l.148-154: Can you explain further, and graphically, what is meant by 'shape, volume, orientation'. EVI, VEV and VVV mean little to me as they currently stand.
We now clarify in the main text: Line 173-175 "The first identifier denotes volume (equal or variable size), the second shape (spherical or not) and the third orientation (aligned or not)."
RESPONSE: See my earlier comments about graphically explaining all this (and all the rest).
See above response

l.158: At this point, I have a very faint idea how your novel approach does what it does. Please help the reader.
We hope that the edits above help with this
RESPONSE: Not much. Please add: (i) a diagram illustrating how the different models/methods are combined; and (ii) a short example demonstrating your novel approach and its advantages 'in action'. See earlier comments.

As above we have added the diagram and further explanation

l.309-322: RESPONSE: Can you please say in the manuscript that you are happy to share your code upon request?
Inserted “Note that all code to reproduce the data analysis is available on request from the authors.” Towards the start of the Data Analysis section

l.464: ‘the’ is missing.
Fixed