

Dear Recommender;

Please find here enclosed our revised version of the manuscript “Lactation curve model with explicit representation of perturbations as a phenotyping tool for dairy livestock precision farming” for recommendation in Peer Community in Animal Science. We are grateful to the reviewers for their input which we feel has helped us to improve the readability and quality of our manuscript.

As suggested by the reviewers, we revised our manuscript to improve the readability of the model description and focus more on ideas and concepts. As shown in the PCI submission process, we submitted a revised version in bioRxiv (<https://www.biorxiv.org/content/10.1101/661249v3>). Hereafter, we described the main modifications done to address these issues.

Revision Round #1

Lactation curve model with explicit representation of perturbations as a phenotyping tool for dairy livestock precision farming.

Reviews:

Review1: by Alberto Atzori

Review of the Manuscript Lactation curve model with explicit representation of perturbations as a phenotyping 1 tool for dairy livestock precision farming.

General comments The manuscript presents a new approach to model the lactation curve using daily records. The approach is highly innovative and offers excellent application opportunities, both for scientific purposes and for managerial strategies at farm level. The model was developed on goat milk records but s highly transferable to other species raised in farms with technological equipment’s allowing to get daily milk records at individual level. The manuscript is clearly written, easily understandable and readable. The mathematical annotation is well integrated within the paragraphs.

The work is very inspiring and valuable. Following the manuscript structure, only minor comments and suggestions, and editorial changes, were reported below:

Minor points:

Line 28: it is not clear the use of the word “confronted”.

We changed the word “confronted” by “face”

Line 31: this effects is perhaps too far from the related noon. A possible modification of the sentence could be: “[..]can be used to evaluate perturbations...”

Done

Line 58: Please check the properness of the term “problems” in respect to milk yield

We changed the word “problems” by “events”

Line 59: cow comfort, welfare and social interaction, as well meteorological changes (not only heat) and feed quality are often related with short term perturbations of lactation curve.

Done

Lines from 85 to 93: these aspects are very interesting for model applications and represent a very good frontier in milk yield prediction. I guess that the presented model could contribute to reach these objective in the future. However the use of milk data and events “as proxy” and “alerts” is not covered by the presented analysis, which is mainly based on data fitting. Thus these arguments should be moved in discussions as possible further extensions of the work, etc. Objective: exhaustive clear and concise.

We moved this paragraph in the discussion part (see lines 545-554)

Line 123: Could be useful to include at least a reference for cattle? Any error was detected in presented equations and math annotation. Many acronyms are reported in the manuscript referring to PLM. A table of acronyms might perhaps help the identification of their meaning without needing to go back and read acronyms in methods.

We added a reference on cattle for Wood (Beever et al., 1991) (see lines 139). We also added a table to help the identification of acronyms referring to PLM and Wood model used in this document (see line132).

Line 300: please check $t_3 \rightarrow 358$ days, whereas in methods max lactation length is declared to be 350 days.

Done

Line 312: please check (0.17 ...), the sentence indicate a decline, thus could it be -0.17?

For more clarity, we added RMSE of Wood model (0.39 ± 0.13 kg) and the RMSE of PLM_N model (0.21 ± 0.05 kg).

Lines 351 to 356: please check if might be important to highlight that atypical curves observed in the dataset could be biologically true (same animal, extended, etc).

The dataset was extracted from the database of our experimental farm where milk records are checked. Therefore, atypical curves are biologically true. In our view, it is important to stress that our model is flexible enough to accommodate atypical production pattern.

Line 347, 348 and 366: A little concern is about the model names. It has to be stated that the paper does not result confusing. Anyway, PLMn seems refer to both unperturbed and perturbed curves. In addition, at line 366 we can read “unperturbed PLM-based model”. It is correct but its explicit name remains “unperturbed perturbed lactation model-based lactation curves”. The use of PLMu could be useful to refer to PLM unperturbed curves even if still contrasting for perturbed-unperturbed. It is understandable that is difficult to find a simple coding for these variables, but being the essential model names it probably deserves a further check.

Done

Lines 406 to 410. Description of parity and age effects is not related with short term perturbations, whereas they are among standard factors affecting lactation curves. Please consider if these sentences are consistent with contents reported from 398 to 403 and 417 to 426.

Indeed, we checked the consistency and move the description of parity and age effects from line 406 to 413 on line 393, accordingly.

Line 442: please consider that the expression of genetic potential of an animal is not the estimated unperturbed curve. The unperturbed curve is also subjected to farm and other environmental conditions in respect to the cow phenotypic potential. Especially if effects with long term consequences influence the farm management. Could the sentence include a specification such as: “...[.genetics] in the studied farm under given environmental constraints” or similar concepts?

Done

From line 446 to 489 the discussion presents a nice synthesis, but focuses on aspects and findings which are more close to introduction and background. Please consider if it can be moved to introduction. Excluding this part from discussions might help to keep focus on the model purposes, use and applications, which are interestingly reported, in the second part of discussions.

We moved this paragraph in the introduction part (see line 83 to 115)

The discussion does not present model limits. From a general point of view, few more explicit examples on possible use and characterization of animal events from precision farming data gathering (such as: feed, rumination, health and reproduction variables or indicators) could help to open a window of model use and development to support managerial practices at farm level (please see comments in introduction for alerts and milk yield early estimates).

As suggested, we added some elements about the model's limits in the discussion (see lines 564-574). This will allow to present model limits and open perspectives for the use of the model.

Tables and Figures:

Table 2. Please consider if decimals have to be reduced according to xxx xx.xx.xx 0.xxx

Done

Typos/ grammatical

Missing spaces were observed at lines: 198, 208, 209, 218, 254, 270, 281, 287.

Done

Missing spaces in lines 316, 386 (missing comma after primiparous), 393, 399, 403 (extraspaces), 419.

Done

Review2: by anonymous reviewer

It is an elegant model in which the power and flexibility of Gamma density functions (and family) highlights their robustness and inferential capabilities of this family of distributions. The model presented an extraordinary adherence to the observed data and the exempli gratia showed as Figure 2 speaks for itself. I dare to say to authors and readers a word of caution about the way Akaike Information Criterion was used. The use of the smallest AIC does not answer the question: how much this model is superior in terms of its content to explain reality? I would like to see the solutions, their respective sums of squares of the residuals, their corrected AIC (it is rather appropriate to sampled data) and their respective model probabilities and evidence ratios. Other equally likely solutions should not be discarded, but averaged with other model solutions together with the smallest AIC solutions if model uncertainty choice is large enough. Another aspect is the treacherous nonlinear algorithm because of possible numerical artifact parameters. Future users of this model have to keep in mind this possibility. I would like to see future developments particularly by solving this model with generalized nonlinear mixed effects tools that otherwise replace the two step analysis of parameter estimates.

Thank you very much for your comments. It is clear that choosing the most reliable criteria in a model is a very delicate topic. We chose Akaike Information Criterion because it is the basic criterion in the nls.multstart algorithm that allows giving a solution of the fit during a run. We repeated this step a one hundred times with more stringent criteria to increase the probability of having the best solution, the best fit. During the development of the model we used several criteria (of course we calculated the sum of squared residuals, and many other methods) and we judged that the AIC offers a satisfactory solution. We also provide the R code so that users of this model can improve it and choose the parameters that fit their needs. We clarified line 510- 514 that the choice of criteria could be user defined in the fitting algorithm.

Review3: by Jennifer Spencer

The manuscript entitled “Lactation curve model with explicit representation of perturbations as a phenotyping tool for livestock precision farming” by Ahmed et al. which describes a model to assess perturbations effects on the lactation curve in dairy goats. This manuscript provides a logical reason for the development of a model to assess the lactation curve

especially during times of perturbations allowing to determine the effects these events have on the lactation curve. Additionally, there appears to be a lack of information on effects of stress on dairy goat lactations, and therefore, this manuscript provides a tool that can help to assess these effects and potentially be used to develop management practices and strategies to improve milk production on dairy operations. This is a well written manuscript and provides a depth of knowledge, however, this manuscript needs some work in terms of organization and flow throughout the paper. Below I have provided a few general suggestions and within the PDF file more specific edits for the authors. These suggestions are mostly editorial and clarification in order to strengthen the article.

General Comments:

Possible keyword may be perturbations since that is what the manuscript is about or even “challenges”.

We added “perturbations” in keyword

Were all goats used for analysis fall kidding?

In our experimental facilities, kidding period is in January and February.

In the discussion explain the parameters for clarity.

As suggesting by reviewer 1, all model names and acronyms of PLM were revised.

In the discussion, could the effects of perturbation 1 have carry over effects onto perturbation 2?

In the PLM model, perturbation parameters are independent so they can be nested. This induces a cumulative effect on the perturbed curve.

Would that be something that the authors would further analyze or investigate?

Indeed, PLM could be used to investigate the biological basis of carry over effects from a perturbation to the next.

Minor points:

Line 19: State as authors did in the Acknowledge section.

Done

Line 46: Possibly another word for disturbing, challenging may be a better choice.

Done

Line 49: loss?

The model allows the characterisation of both potential milk curve and loss induced by perturbations.

Line 51: what kind?

Given that the numbers of words in the abstract are limited, we changed "decision solutions" by "decision support tools" for clarification without details.

Line 83: What do the authors mean by resilience?

As reviewer 1 mentioned, introduction should be concise and clear. Therefore, as resilience is now commonly used in animal sciences literature, we did not develop definitions but we added a relevant reference

Line 85: Please add something like "develop best management strategies to maximize performance". Adds a little more information.

We added "Evaluating the effect of perturbations on animal performance could provide metrics to quantify how animals cope with their environment, and develop best management strategies to find an optimal balance between animal welfare and performance."

Line 92: Are the authors sure that is it per hour or per milking? Please clarify.

We checked in the original publication and it's per hour.

Line 98: What do the authors mean by smoothing techniques?

We moved "such as smoothing methods," line 91 to link with the reference of Codrea et al. that provide information on these methods Globally, smoothing techniques consist in reducing or limiting the discontinuities of the first / second derivative.

Line 100: Would this be the quantification of perturbation effects on money or continued performance or other factors?

We added "animal performance"

Line 132: What is "e" in the equation? I assume it is error.

"e" is the mathematical symbol for the exponential function

Line 158: Flip] to [.

We chose to put a close bracket because the zero value is excluded from the interval of value that the parameter can take ($k_{0,i}$ never equals zero)

Line 192: if(t) may explain better if t seems like its own variable not if the time

It's time, we added a space.

Line 256: Abbreviate here (RMSE) as it is used later in the manuscript.

Done

Line 273: Fall born, were breeds evenly distributed between the years? Also for the data from the 319 and then this data with 181 goats the authors may consider calling each one experiment 1 and experiment 2 or something to that extent.

The numbers of animals per breeds were more or less evenly distributed between birth years. The data set of 319 lactations comes from 181 goats, we do not have two experiments. We rephrased in manuscript (see lines 288-294) for more clarity

Line 273: Is it normal for a goat to lactate this long? Why did the authors choose this day?

Yes it is. We did not choose this day, it is the maximum number of days in milk in our dataset

Line 280-284: Please reword, confusing.

We changed “Fixed effects of breed (Saanen vs. Alpine) and parity (1 vs.2 and more) were tested on RMSE, on parameters a, b, and c, on estimated peak milk yield, peak time and total milk yield over $[t_0; t_3]$ for Wood and PLM models, and on the estimated number of perturbation and percentage loss for PLM model, with a mixed analysis of variance model with goat as a random factor” by “Fixed effects of breed (Saanen vs. Alpine) and parity (1 vs.2 and more) were tested on parameters of Wood, with and without the changes made from PLM model. It was also tested on estimated peak milk yield, peak time, total milk yield over $[t_0; t_3]$, the number of perturbation and the rate milk loss using a mixed analysis of variance model with goat as a random factor.”

Line 298: Of milk? Please clarify.

We added milk for clarification.

Line 300: Are brackets needed? I am not quite sure how these values are lining up in the sentence. May need to explain 4 was ... and 11 was ... for clarification.

We replace “the record interval length $t_0 - t_3$ was between $[1, 5]$ to $[165, 358]$ in 300 days)” by “the record interval length was between 1 and 5 days for t_0 and between 165 and 358 days for t_3 .”

Line 309: Time to peak or time of peak?

It was time of peak; we added it in the text.

Line 327: This is referring to wood's model or is it for both models? Please state.

*We added “(for both models: PLM^{*N} and Wood model)”*

Line 329: I am not too sure how the authors are getting this number could they explain the math

We added in the Table 1 the origins of these values. We also separated the values according to the breeds

Line 337: Again review is confused where these values are at on the table.

We added in the Table 2 the origins of these values.

Line 341: What are the units for these parameter values? Same with the table.

The units of Wood's parameters are present in the material and method section (line 145-151)

Line 341: These are very small numbers so does this have any significance in actual practice?

The parameter c in the wood equation presented line 357 is multiplied by time in an exponential function. As a result, a small variation in the value of this parameter caused a great change in the result.

Line 356: 18 curves out of how many curves?

We added “A total of 18 out of the 319 analysed curves...”

Line 364: This last sentence seems to fit better in the discussion section.

We kept this sentence in line 381

Line 395: But rather... Maybe a combination of things such as and could list.

In this part, we do not wish to present all hypothesis around these results, the main aim here is primarily to present PLM and the kind of result that we can extract with PLM from raw data.
We can't conjecture on other factors affecting milk loss in this paragraph. The result we obtained with the model did not allow us to make speculations on milk loss.

Line 402: Put brackets around k_0 and state "The intensity (k_0) ...

We explained the parameter k_0 when first used in the manuscript. It stands for intensity of a perturbation.

Line 403: What do the authors mean by intense, worsen or have more of an effect, could explain a little more or give an example of what they observed.

We added "(Table 3)" because here we are talking about the k_0 parameter (which represents the intensity of the perturbation) and we only present the results obtained in Table 3

Line 418-422: This section is a little confusing. Can the authors please clarify what they are meaning by the collapse speed, speed recovery, etc. Possibly give an example of when these would occur in a lactation.

All parameters are explicitly defined in Material and methods section. To avoid redundancy, we did not repeat definitions throughout manuscript. Yet, as proposed by reviewer 1, we added a compendium table (Table1) summarizing all model components.

Line 438: models used to assess perturbations during a lactation in livestock?

Wood's model does not assess perturbations during lactation in livestock but gives us an idea on milk yield during lactation without taking account of perturbations.

We kept initial phrasing: the Wood model was made to assess the general pattern of lactation curve and how this curve can be affected by factors (parity, ...). This model was not made to assess perturbations that affect the general pattern.

Line 438: This could be moved to the very beginning of this discussion section. Also add the after therefore, and delete of after fitting. Also give examples of external factors that alter the lactation curve for clarity.

We moved this section in introduction, as also suggested by reviewer 1 (background elements). We kept the discussion focused on the model itself and not on lactation.

Line 458: Which is decline in production?

Not necessarily, it is to take account of variations that may exist in the lactation data at a time t and at $t + 1$

Line 479: Of the lactation curve, of milk production?

We added "of the lactation curve and intensity of milk production"

Line 482: What are the authors referring to when they state "objects"?

We added: "cited above"

Line 486: Which is a challenge or stress.

We replace "event" by "stress"

Line 489: In what species ?
We added: "growing pigs"

Line 535 : identifying what?
We added "parameters"

Line 535: explain what this parameter is
We added: (intensity of the perturbation)

Line 558: milk production or general production
We added "milk"

Line 572: Not to sure what the authors mean by diet quality perhaps feed availability?
We replaced "quality" by "feed availability"

Line 595: Compare animals phenotypic potential?
We added "animals' phenotypic potential"

Line 599: What do the authors mean by biological framework?
Development of models applied to animal sciences requires to keep in mind that we are working on living organisms. Therefore, we aim at using concepts and mathematical tools without losing the connection with biological objects. In a pure mathematical view, parameters do not necessarily make sense for biology. With a biological framework, the parameters coming from the maths still have meaning for biology.

Typos/ grammatical:

Line 30, 59, 69, 77, 157, 188, 214, 265, 291, 360,363,389, 397, 407 and 502: Insert comma
Done

Line 116, 130, 131, 135, 151, 198, 200, 208, 218, 219, 222, 246, 247, 248, 253, 254, 270, 290, 393 and 399: Insert space.
Done

Line 20: Capitalize "I"
Done

Line 22: Insert ":"
Done

Line 27: add "effects"
Done

Line 39: add "for"
Done

Line 48: add "for"
Done

Line 61: Change “in” by “of”.

Done

Line 63: Change “mammal females” by “mammalian females,”

Done

Line 66: at peak milk yield?

We added “maximum”

Line 69: "of the" milk yield and in preparation for parturition.

Done

Line 70: Please delete

Done

Line 74: Delete and insert published by Wood in 1976.

We deleted.

Line 75: Please delete

Done

Line 76: replace “extracting” by “creating”

Done

Line 76: replace “and therefore” by “thereby”

Done

Line 77: tool instead of tools

Done

Line 80: Insert “the”

Done

Line 87: Replace “[11]. Studying” by “[11], studying”.

We kept the two sentences separated as they address different topics.

Line 91: Change milk anomalies for milk abnormalities.

Done

Line 96: change “we can move from the logic of reducing variability around average profiles to the logic of extracting variability to provide information as such” by “we can transition data analysis from reducing variability around average profiles to extracting variability to provide information..such as”

Done

Line 97: New paragraph.

Done

Line 98: Please delete only need the number that corresponds to the reference

Done

Line 99: Please italicize

Done

Line 103: Change “There is thus” by “Thus, there is”

Done

Line 181: were instead of are

Done

Line 181: no comma

Done

Line 193: Add an s to end

Done

Line 250: change “This step” by “Step 1”.

Done

Line 265: italicize

Done

Line 291-292: Do you need brackets for these values or could you state 1.21 to 34.45, 34.45 to 171.05, and 171.05 to 270.30, respectively? Please add days before respectively as well.

Done

Line 294: Delete “and”

Done

Line 295: Add per animal per lactation.

Done

Line 306: Add an "s" to perturbation.

Done

Line 334: Insert (Table 1)

Done

Line 385: Delete brackets. State were not significant and were ...

Done

Line 387: Insert after perturbation.

Done

Line 430: Delete 1. Go through and make sure that headings and subheadings are consistent throughout the paper.

Done

Line 437: change “illustring” by “Illustration of”

Done

Line 437: Affected instead of affect.

Done

Line 460: to assess?

Done

Line 467: Add during the lactation

Done

Line 469: Add an s to challenge and put dairy cows

Done

Line 470: Insert an

Done

Line 471: Replace with study

Done

Line 472: Perhaps feed restriction during instead of off-feed since it is somewhat jargon.

Done

Line 473: Insert On the otherhand,.

Done

Line 475: Only one other study referenced

We replaced by “Another study”

Line 476: Diet instead of dietary

Done

Line 477: With instead of to

Done

Line 478: Insert in

Done

Line 488: The instead of to.

Done

Line 488: Characterization of

Done

Line 493: spelling

Done

Line 495: during the lactation curve

We added “during lactation”

Line 495: Another word for controlled.

We kept the word "controlled" and we reformulated the sentence in the manuscript

Line 501: lactating female. Delete which.

Done

Line 502: disease as well?

We added "disease"

Line 502: Sentence as "data such as body weight changes, dry matter intake, and hormones differing during lactation?"

We added "With the development of on-farm technology measurements, an interesting perspective for the PLM model is to be used to assess other biological time-series data, such as body weight changes, dry matter intake, and hormones differing during lactation."

Line 518: Substantially different?

We added "Substantially different"

Line 519: delete

Done

Line 520: Insert during lactation

Done

Line 522: was instead of is

Done

Line 523: was instead of is

Done

Line 524: State "difficulty of the PLM model"

Done

Line 524: was instead of has been

Done

Line 527: delete "human hand"

Done

Line 528: Delete "could be the curve" insert the lactation curve would be

Done

Line 529: Delete "the human eye"

Done

Line 529: Insert been.

Done

Line 529: Delete "although" insert yet.

Done

Line 530: Insert PLM.

Done

Line 532: insert the word estimates.

Done

Line 533: replace “shown” by “showed”.

Done

Line 535: replace “shown” by “showed”.

Done

Line 538: Insert comma and then delete that..

Done

Line 549: Delete brackets. Put period and capitalize For

Done

Line 551: parameters, or .period.

Done

Line 556: Spell out, "PLM was developed"

Done

Line 556: The instead of our

Done

Line 560: Insert assuming no perturbations were encountered.

Done

Line 561: replace "to cull" by "the culling of".

Done

Line 562: replace "and decreased" by "decreasing".

Done

Line 562: Insert milk.

Done

Line 562: Insert yet instead of while.

Done

Line 569: animals, such as...., and can

Done

Line 575: on the instead

Done

Line 575: delete "scale"

Done

Line 575: Made instead

Done

Line 576: Insert understanding.

Done

Line 577: replace "is crucial" by "during critical times in order" ...

Done

Line 581: Replace "(e.g)" with "by" ... Delete and use "and"

Done

Line 586: Replace "(intensity, collapse, recovery)" by "Of each and subsequent perturbation."

Done

Line 586: Replace "it" with The PLM model

Done

Line 587: Add s in "responses" and "due to specific perturbations."

Done

Line 588: insert can

Done

Line 589: insert and selection

Done

Line 592: Add effects on

Done

Line 593: Replace "reflecting" with "allowing to assess"

Done

Line 596: replace "is" with "could be used" as a since the authors have not shown this specifically

Done

No spaces between number and % symbol (2 % to 2%).

Done

For numbers such as 10000 please use commas 10,000 in manuscript and tables.

Done

Consistency for headings and subheadings (i.e. headings use capital for each major word and bold; subheadings bold and italicized).

Done

Tables and Figures:

Figure 1, Line 143: Is the x-axis referring to the mean milk being drawn from the udder both sides or just one quarter?

The x-axis refers to time, expressed in days after perturbation. The y-axis refers to the proportion of milk that is withdrawn from the potential milk production because of the perturbation.

Line 145: Insert space.

Done

Figure 2, Line 143: Adding more space between figure 1a and figure 1b may help the reader clearly identify each part.

Done

Figure 3, Line 316: Insert space.

Done

Line 317: State that this is for one animal (*i.e.* an example or representative of one animal's perturbed lactation model).

We clarified the caption of the figure

Table 1, line 319: Please use commas for values 1,000 and above. Also please round values to the nearest 10th (two values after the decimal).

We added the commas for values 1,000 and above.

Line 319: Please add more details about what this table is about. For example, comparison between breeds and lactation numbers across the models and variables.

We added “: comparison between breeds and lactation numbers across the models and variables”

Line 319: Please capitalize peak here and throughout this column.

Done

Line 320: P = or \leq these values to have these codes?

We added \leq

Line 322-323-224: Insert space

Done

Figure 4, Line 366: Delete brackets

Done

Line 367: Capitalize A

Done

Figure 5, Line 380: This figure might need to be a full page so that it is easier to see.

Line 381: Insert the word Panels

Done

Line 382: Delete space after d

Done

Figure 5, Line 553: Period instead of:

Done

Line 553: No space or space between: for all of the parameters.

Done