Package: EMOTIONS (via r-universe)

May 31, 2026

Type Package

Title Ensemble Models for Lactation Curves

Version 1.3

Description Lactation curves describe temporal changes in milk yield and are key to breeding and managing dairy animals more efficiently. The use of ensemble modeling, which consists of combining predictions from multiple models, has the potential to yields more accurate and robust estimates of lactation patterns than relying solely on single model estimates. The package EMOTIONS fits 47 models for lactation curves and creates ensemble models using model averaging based on Akaike information criterion (AIC), Bayesian information criterion (BIC), root mean square percentage error (RMSPE) and mean squared error (MAE), variance of the predictions, cosine similarity for each model's predictions, and Bayesian Model Average (BMA). The daily production values predicted through the ensemble models can be used to estimate resilience indicators in the package. The package allows the graphical visualization of the model ranks and the predicted lactation curves. Additionally, the packages allows the user to detect milk loss events and estimate residual-based resilience indicators.

Depends R (>= 4.2)

Imports dplyr, orthopolynom, quantreg, minpack.lm, tidyr, ggplot2, ggridges, parameters, rlang, tidyselect, splines, tibble

License GPL-3

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Suggests knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Pablo Fonseca [aut, cre], Marcos Prates [aut], Aroa Suarez-Vega [aut], Ruth Arribas Gonzalo [aut], Beatriz Gutierrez-Gil [aut], Juan José Arranz [aut]

Maintainer Pablo Fonseca <p.fonseca@csic.es>

Repository https://cran.r-universe.dev

Date/Publication 2026-04-01 21:04:06 UTC

RemoteUrl https://github.com/cran/EMOTIONS

RemoteRef HEAD

RemoteSha 9b5b5dd3456f1f00dc7fd61893969e26e2972f71

Contents

`imp_my` *Impute missing daily milk yields using the ensemble created*

Description

Impute missing daily milk yields using the ensemble created

Usage

```
imp_my(out.ind = NULL, dim = NULL)
```

Arguments

`out.ind`	The list containing the data frames with the daily production records obtained from the LacCurveFit function
`dim`	A vector with the days in milk where the milk yield will be imputed. It can contain observed and missing DIM

Value

A data frame containing the imputed milk yields for the days in milk informed.

`LacCurveFit` *A wrap function to the ModelsLac function that allows the fit of lactation curve models based on daily production and days in milk records simultaneously for a list of animals*

Description

The function uses a data frame containing the daily milking records as input

Usage

```
LacCurveFit(
  data,
  ID,
  trait,
  dim,
  alpha = 0.5,
  models = "All",
  param_list = NULL,
  silent = T
)
```

Arguments

`data`	A data frame containing the daily milking records
`ID`	The name of the column containing the unique IDs of the individuals
`trait`	The name of the column containing daily milking records
`dim`	The name of the column containing days in milk records
`alpha`	A penalization factor, ranging from 0 to 1, for the estimation of the model's weight
`models`	A vector describing the models to be included in the analysis. In total, 47 models are included in EMOTIONS. The default option is "All", which results in the inclusion of the 47 models. Alternatively, a vector containing any subset of the following models can be provided: "MMR","MME","brody23","brody24", "SCH","SCHL","PBE","wood","DHA", "CB","QP","CLD","PapBo1","PapBo2", "PapBo3", "PapBo4", "PapBo6", "GS1", "GS2","LQ", "wil", "wilk", "wilycsml", "BC", "DJK","MG2", "MG4", "MG", "KHN", "AS", "FRP","PTmult","PTmod", "MonoG", "MonoGpw", "DiG", "DiGpw","legpol3", "legpol4", "legpolWil", "cubsplin3", "cubsplin4", "cubsplin5", "cubsplindef", "wilminkPop", "qntReg".
`param_list`	A list composed by the models, named as in the models parameter, and the repective parameters included in the models.
`silent`	A logical string defining if warning should be printed or not during the model fitting. The default value is TRUE (not printing warnings).

Value

A list containing the fitted models, the model‘s weigths and ranks for each weighting strategy, and the predicted daily production obtained through the model ensemble for each weighting strategy

---

`milkloss_detect` *Identify milk loss events and resilience indicators from daily milk yields*

Description

Identify milk loss events and resilience indicators from daily milk yields

Usage

```
milkloss_detect(
  data,
  id_col,
  dim_col,
  MY_col = "MY_real",
  MY_pred,
  dim_start = 1L,
  dim_end = 305L,
  rec_mode = c("pctbase", "band", "resid"),
  drop_pct = 0.1,
  min_len = 1L,
  tol = 0.05,
  stick = 3L,
  rec = 1
)
```

Arguments

`data`	A data frame containing the observed and predicted daily milking records.
`id_col`	The name of the column containing the individual IDs.
`dim_col`	The name of the column containing the days in milk.
`MY_col`	The name of the column containing the observed milk yield.
`MY_pred`	The name of the column containing the predicted milk yield (baseline).
`dim_start`	The first day in milk to consider when identifying milk loss events and resilience indicators.
`dim_end`	The last day in milk to consider when identifying milk loss events and resilience indicators.

rec_mode	How "recovery" is defined. One of: "pctbase": recovery when the observed value reaches a given fraction of the baseline (rec), for a given number of consecutive days (stick); "band": recovery when the observation is inside a tolerance band around the baseline (+/- tol), for at least stick consecutive days; "resid": recovery when the residual has improved enough from the nadir (by a fraction rec of the nadir's absolute residual) for stick consecutive days.
drop_pct	Minimum relative drop from the anchor (baseline reference) to accept an episode.
min_len	Minimum number of consecutive days with negative residuals required to define an episode.
tol	Used when the "band" mode is selected. Half-width of the tolerance band around baseline in relative terms.
stick	Maximum number of consecutive days in recovery to consider an episode finished.
rec	Minimum relative recovery from the nadir to finish an episode (used in "pctbase" and "resid" modes).

Details

The function computes several descriptors of milk-yield perturbation episodes.

1) Nadir (day of minimum)

The worst day inside the episode (deepest point of the perturbation).

`t_hat = argmin_{t in [t_start, t_end]} obs(t)`

`Nadir = obs(t_hat)`

where `t_start` and `t_end` are the episode boundaries.

2) Amplitude (drop)

Depth of the dip relative to the baseline at the episode start.

`A = baseline(t_start) - obs(t_hat)`

Some variants use `baseline(t_hat)` instead of `baseline(t_start)`; here the start of the episode is used as the reference.

3) ML_per_event (AUD)

Total milk lost (in baseline units) over the episode, i.e., the integrated milk deficit.

`ML_per_event = AUD = sum_{t=t_start..t_end} [baseline(t) - obs(t)]`

In discrete data, AUD is computed with day-weighting: each observation contributes

`(baseline(t) - obs(t)) * delta_days`

where `delta_days` is the gap to the next observed DIM (last day weight = 1).

4) Time-to-baseline (TTB)

Time after the nadir until the profile returns to (and stays near) the baseline.

Recovery is declared when `obs(t)` re-enters a tolerance band around the baseline and stays there for `stick` consecutive days (controlled by `tol` and `stick`).

We find the smallest `tau >= 0` such that for all `u` in the interval from `t_hat + tau` to `t_hat + tau + stick - 1`:

`abs(obs(u) - baseline(u)) <= tol * baseline(u)`

Then: `TTB = tau`.

If this condition is never satisfied before DIM 305, `TTB` is set to `NA` (right-censored).

5) Recovery half-life (t_1_2)

Earliest time after nadir when half of the drop has been recovered.

With amplitude `A` as above, define the half-recovery level:

`L_half = baseline(t_start) - A / 2`

Then:

`t_1_2 = min{tau >= 0 : obs(t_hat + tau) >= L_half}`

6) Slopes (decline and recovery)

Average daily change during the decline into the nadir and during early recovery, summarizing the episode shape.

For a `K`-day local window:

`DeclineSlope = (obs(min(t_hat, t_start + K)) - obs(t_start)) / (min(t_hat, t_start + K) - t_start)`

`RecoverySlope = (obs(min(t_end, t_hat + K)) - obs(t_hat)) / (min(t_end, t_hat + K) - t_hat)`

7) AUC_deviation

Trapezoidal area under the curve of the milk deficit `baseline(t) - obs(t)` across the whole episode. It summarizes how much milk was lost and for how long.

Conceptually:

`AUC_deviation = integral_{t_start..t_end} [baseline(t) - obs(t)] dt`

In practice this is approximated via the trapezoidal rule on discrete DIMs.

8) prod_decline_slope_amp

Product of the decline slope (anchor -> nadir) and the amplitude (anchor - nadir). It combines speed and depth of the decline into a single indicator of how "aggressive" the drop is.

`prod_decline_slope_amp = DeclineSlope * A`

9) prod_recovery_slope_TTB

Product of the recovery slope (nadir -> recovery) and time-to-baseline (TTB). It combines how fast the animal recovers with how long recovery takes, summarizing recovery efficiency.

`prod_recovery_slope_TTB = RecoverySlope * TTB`

Value

A list with two data frames:

- `episodes`: individual milk loss events and their resilience indicators;
- `aggregates`: milk loss events aggregated per individual.

The resilience indicators identified are described in the Details section.

`ModelRankRange` *Create a line plot that shows the range of the ranks obtained for each model across the individuals*

Description

Create a line plot that shows the range of the ranks obtained for each model across the individuals

Usage

```
ModelRankRange(LacCurveFit, metric = "AIC_rank")
```

Arguments

`LacCurveFit`	The object obtained from the LacCurveFit function
`metric`	The name of the metric to be use to plot the model´s ranks

Value

A line plot that shows the range of the ranks obtained for each model across the individuals

`ModelsLac` *Performs the model fitting and the weight assignment based on different strategies for each individual ID*

Description

Performs the model fitting and the weight assignment based on different strategies for each individual ID

Usage

```
ModelsLac(
  data,
  ID_col,
  ID,
  trait,
  dim,
  alpha,
  models,
  param_list = NULL,
  silent = silent
)
```

Arguments

data	A data frame containing the daily milking records
ID_col	The name of the column containing the unique IDs of the individuals
ID	The individual ID that is being analyzed
trait	The name of the column containing daily milking records
dim	The name of the column containing days in milk records
alpha	A penalization factor, ranging from 0 to 1, for the estimation of the model's weight
models	A vector describing the models to be included in the analysis. In total, 47 models are included in EMOTIONS. The default option is "All", which results in the inclusion of the 47 models. Alternatively, a vector containing any subset of the following models can be provided: "MMR","MME","brody23","brody24", "SCH","SCHL","PBE","wood","DHA", "CB","QP","CLD","PapBo1","PapBo2", "PapBo3", "PapBo4", "PapBo6", "GS1", "GS2","LQ", "wil", "wilk", "wilycsml", "BC", "DJK","MG2", "MG4", "MG", "KHN", "AS", "FRP","PTmult","PTmod", "MonoG", "MonoGpw", "DiG", "DiGpw","legpol3", "legpol4", "legpolWil", "cubsplin3", "cubsplin4", "cubsplin5", "cubsplindef", "wilminkPop", "qntReg"
param_list	A list composed by the models, named as in the models parameter, and the repective parameters included in the models.
silent	A logical string defining if warning should be printed or not during the model fitting. The defaul is TRUE (not printing warnings).

Value

A list containing the fitted models, the model's weigths and ranks, and the predicted daily production obtained through the model ensemble

---

PlotMilkLoss *Plot the actual daily milk daily production and the predicted values highlighting the detected milk loss events*

Description

Plot the actual daily milk daily production and the predicted values highlighting the detected milk loss events

Usage

```
PlotMilkLoss(
  data,
  ID,
  res.milkloss,
  MY_col,
  MY_pred,
  col = c("red", "blue", "darkgreen"),
  id_col = "ID"
)
```

Arguments

data	A data frame containing the observed and predicted daily milking records
ID	The ID of the individual that will have the daily milking records plotted
res.milkloss	The object with the output of milkloss_detect function
MY_col	The name of the column containing the observed milk yield
MY_pred	The name of the column containing the predicted milk yield
col	The colors of the actual, predicted values, and milk loss events. In this order
id_col	The name of the column containing the individual IDs

Value

A plot with the actual daily milk daily production and the predicted values highlighting the detected milk loss events

---

PlotWeightLac *Plot the actual daily milk daily production and the predicted values obtained by the ensemble model*

Description

Plot the actual daily milk daily production and the predicted values obtained by the ensemble model

Usage

```
PlotWeightLac(
  data,
  ID,
  trait,
  metric,
  dim,
  col = c("red", "blue"),
  point_size = 2,
  line_size = 1,
  axis_text_size = 15,
  axis_title_size = 15
)
```

Arguments

data	The object generated by the LacCurveFit.
ID	The ID of the individual that will have the daily milking records plotted
trait	The name of the column containing daily milking records.
metric	The name of the strategy used obtained the predicted values through the ensemble model.

dim	The name of the column containing days in milk records
col	The colors of the actual and predicted values.
point_size	Numeric value indicating the size of the observed data points in the plot.
line_size	Numeric value indicating the thickness of the regression line.
axis_text_size	Numeric value defining the font size of the axis tick labels.
axis_title_size	Numeric value defining the font size of the axis titles.

Value

A plot with the actual and predicted daily milk production across the days in milk

---

ResInd *A function to estimate resilience estimators (logarithm of variance, lag1 autocorrelation and skewness) based on daily milk production records*

Description

A function to estimate resilience estimators (logarithm of variance, lag1 autocorrelation and skewness) based on daily milk production records

Usage

```
ResInd(
  production_df,
  dim_filter_range = c(1, 7, 203, 210),
  outlier_sd_threshold = 4,
  weight = "weight_AIC",
  trait,
  DIM,
  ID_col
)
```

Arguments

production_df	The list containing the data frames with the daily production records (actual or predicted) obtained from the LacCurveFit function
dim_filter_range	A vector containing the lower and upper limits to remove lactation records from the begin and end of the lactation, if needed. If it is not necessary to remove daily records, the first two values can be set as the minimum days in milk value and the last two as the maximum days in milk values
outlier_sd_threshold	A threshold defining the maximum standard deviations to consider an individual resilience indicator value

weight	The name of the column containing the selected ensemble prediction. The default is weight_AIC
trait	The name of the column containing daily milking records
DIM	The name of the column containing days in milk records
ID_col	The name of the column containing the unique IDs of the individuals

Value

A list containing the daily milk production values after filtering, the list of removed animals, and a data frame with the resilience indicators

---

`RidgeModels` *The function RidgeModels allows the visualization of the distribution of model's ranks across individuals using ridge density plots*

Description

The function RidgeModels allows the visualization of the distribution of model's ranks across individuals using ridge density plots

Usage

```
RidgeModels(LacCurveFit, metric = "AIC_rank")
```

Arguments

LacCurveFit	The object obtained from the LacCurveFit function
metric	The name of the metric to be use to plot the model´s ranks

Value

A ridge density plots for the models included in the ensemble

Index