nature metabolism

Supplementary information

https://doi.org/10.1038/s42255-025-01316-8

Calopy — an advanced framework for the integration and analysis of indirect calorimetry data

In the format provided by the authors and unedited

An Advanced Framework for the Integration and Analysis of Indirect Calorimetry Data

https://www.calopy.app

USER HOWTO

Version 1.0
May 6th 2025

Table of Contents

INTRODUCTION

Summary of **Calopy** Software

Calopy is an **open-source, web-based application** for **advanced analysis of indirect calorimetry (IC) data**, developed using **Shiny for Python**. It is accessible both **online and locally**, offering a **transparent and flexible** platform for researchers.

Calopy integrates **advanced preprocessing capabilities**, including **outlier detection, subject removal, and time-resolved filtering**, allowing researchers to extract key metabolic features like **maxima, minima, amplitude, and area under the curve (AUC)**. It also introduces an innovative **estimation method for resting metabolic rate (RMR) and basal metabolic rate (BMR)** using regression models that incorporate **activity and food intake**.

For data analysis, **Calopy** provides a **comprehensive statistical toolbox**, supporting tests on both **raw data and derived features**, with methods including **Between-Group Comparison (BGP), Temporal-Condition Analysis, and Time-Window Comparison**. All analyses are **non-destructive**, meaning data can be adjusted or reprocessed at any time.

With its **intuitive interface, reproducible results, and open-source framework**, **Calopy** enables researchers to **perform, extend, and refine** IC data analysis with ease, filling a crucial gap in the field.

Calopy is designed for an intuitive and exploratory analysis of Indirect Calorimetry (IC) data. The core of **Calopy** is based on its strict separation of data handling and data analysis.

Summary of **Calopy** Data Types

Calopy handles two main data types: **Metabolic Variables (MVs)** and **Metadata**.

- **Metabolic Variables (MVs)** include time-series measurements such as **VO_2, VCO_2, EE, and RER**, as well as additional data like **food intake, water intake, and activity**. Calopy supports **feature extraction**, including **mean, amplitude, area under the curve (AUC), and peak detection**.
- **Metadata** provides contextual information about each subject and is divided into **Phenotypic Variables (PVs)** (e.g., body weight, age) and **Conditional Variables (CVs)** (e.g., genotype, treatment). Metadata can be stored within the main data file or uploaded separately.

Including **body weight as metadata** is strongly recommended for **standardized analysis**, and additional metadata can be manually added or imported within **Calopy**.

For additional documentation of our data formats and data types please check our APP documentation.

Supported File Formats for **Calopy**

Currently **Calopy** Supports three main file formats (see **Documentation** for details)

- **Generic**, compatible with **Sable Systems**
- **TSE PhenoMaster**, supports various TSE output formats
- **Columbus Instruments**

Generic CSV files (Sable Systems)

Allow flexible data structuring, supporting both **metabolic data and metadata**, with mandatory columns for **Sample ID, Date/Time, and Metabolic Variables** (e.g., VO_2, VCO_2, EE). Metadata can be included in the same file or separately. Our generic file format is fully compatible with Sable Systems File format.

TSE PhenoMaster

If you're working with **TSE PhenoMaster files**, just upload the **raw, unmodified file**—**Calopy** will handle the rest. These files should have a **property section, a blank separator line, and a measurement table** with consistent time points for accurate analysis.

Columbus Instruments

As CI's software CI-Link exports the data for each subject/cage in a single csv file, for now we support these data only through a small workaround: **Save all single csv files and combine them in an excel file (xlsx)** with one sheet for each cage. Name sheets as '**Cage1', 'Cage2', ...**
This file can be now imported as Columbus Instruments data file. Remember to update your Metadata, by either creating a Metadata file or editing metadata directly. Metadata in the header of the CI exported files are automatically detected and included in the metadata.

NOTE: dependent on your CI Link version data may be exported normalized by mass. As we strongly recommend NOT normalizing your metabolic variables by mass, please check and eventually correct. We are currently working on an automated detection and correction method!

DATA PROCESSING

Preprocessing

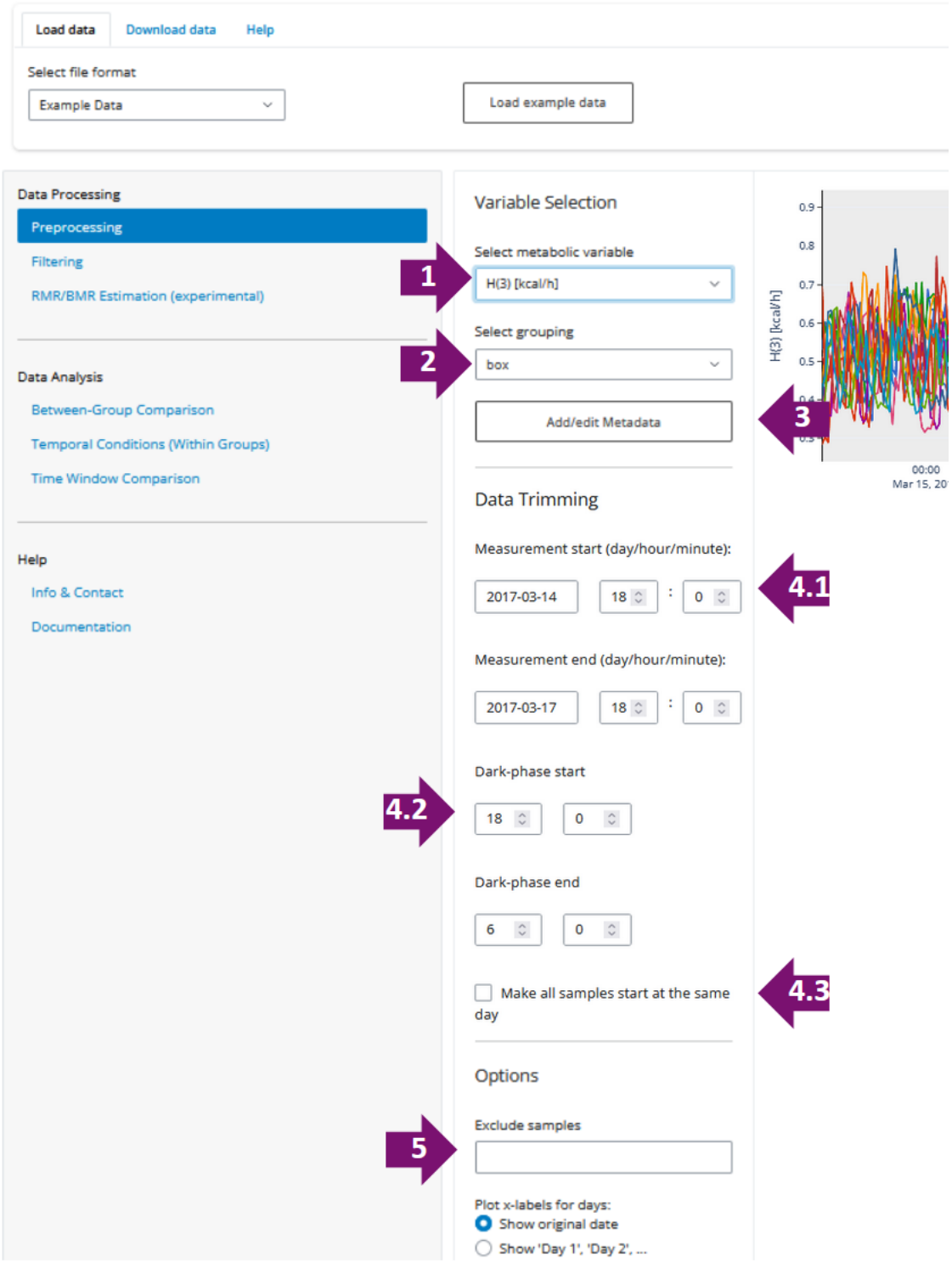


***Figure 1**: Data Preprocessing*

After the data upload you may want to preprocess your data (Figure 1).

1. **Metabolic Variable:** Select Variable to be displayed
2. **grouping:** If groups are defined in your dataset, you may select here for visualization.
3. **Add/edit Metadata:** Opens a dialog where you can upload an additional Metadata file (See file Formats), or directly edit your metadata (Figure 2)
4. **Data trimming:**
   4.1 Measurement start/end: Non-destructive setting of your data start and end time. All the changes you make here define your Experimental start and end time. This is useful to only

include full days for your analysis or remove acclimatization times at the beginning of your experiment.

4.2 dark/light phase: Select your dark /light phases, as indicated in the plot (grey = dark phase). These settings affect your analysis, once you select only dark/light phase for between analysis.

4.3 make samples all start the same day: If your experiment is conducted is multiple batches you can align them to the first day. (Use with care)

5. **Options**:
    - Exclude Samples: Remove individual subjects from your data analysis.
    - Plot x-labels for days: Plot Day-numbers instead of full date.

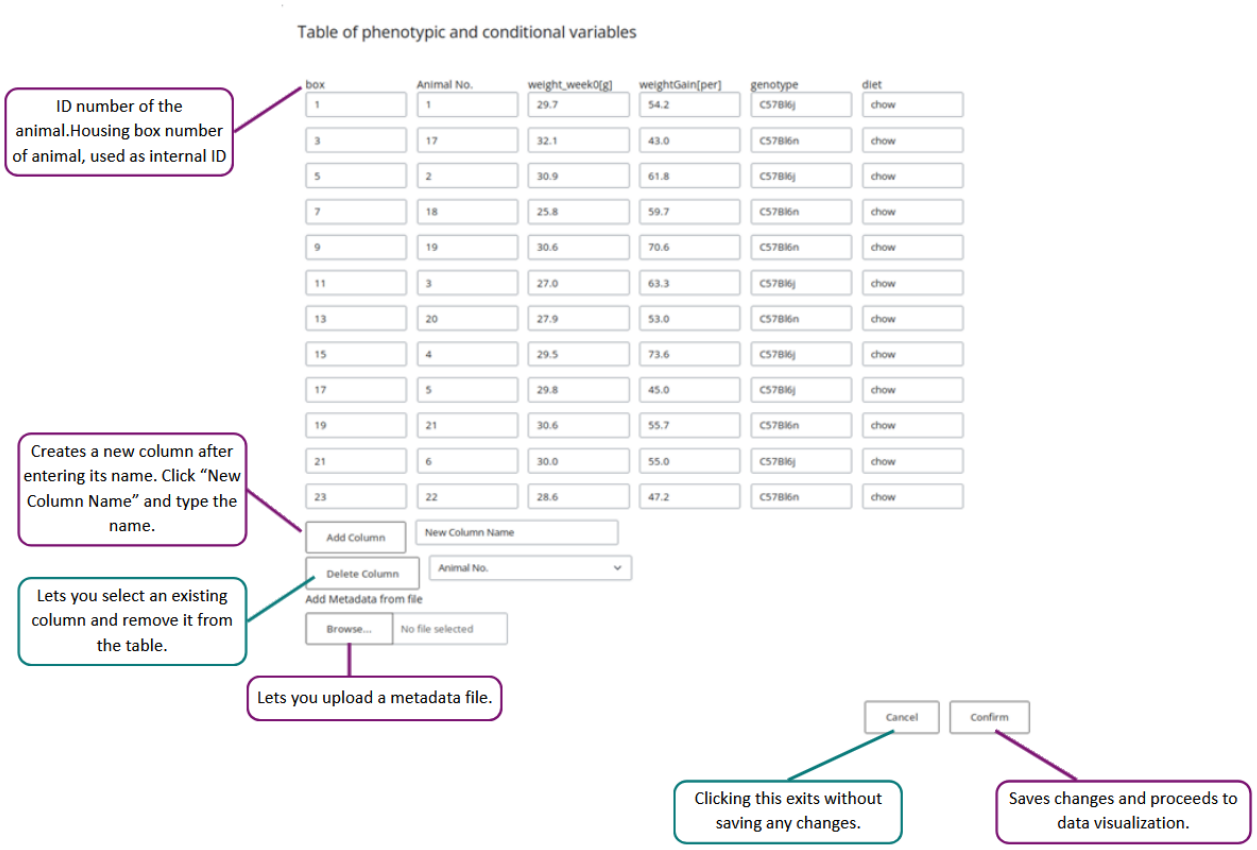


Figure 2: *Upload and edit metadata dialog.*

Filtering

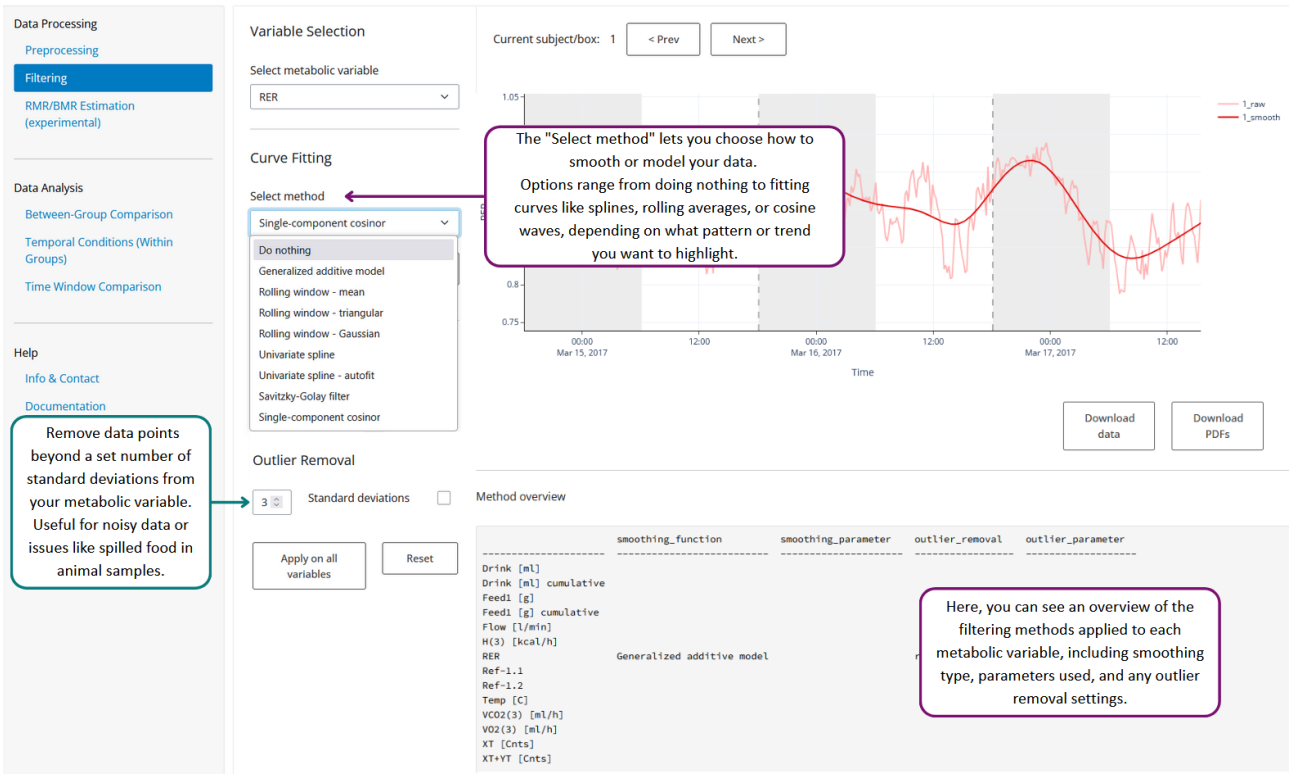


Figure 3: *Data filtering.*

Data filtering allows you to apply smoothing nonlinear filter to your data (curve fitting) or filter for outliers.

Curve fitting allows you to add additional data filters to your data. This may be for various purposes like data smoothing or feature extraction from denoised data.

Smoothing or denoising may be useful for cleaner data plotting and may be used to extract additional useful features (see feature extraction) from your data such as maxima of minima or amplitude.

Use filtering with care since this strongly affects your data and may lead to misinterpretations or erroneous results.

Outlier Removal: Select outlier removal based on x times of the standard deviation of your metabolic variable. This may be useful for very noisy data or e.g., food spilling animal samples.

Both: Curve fitting and Outlier removal can be individually applied to different metabolic variables.

RMR/BMR Estimation

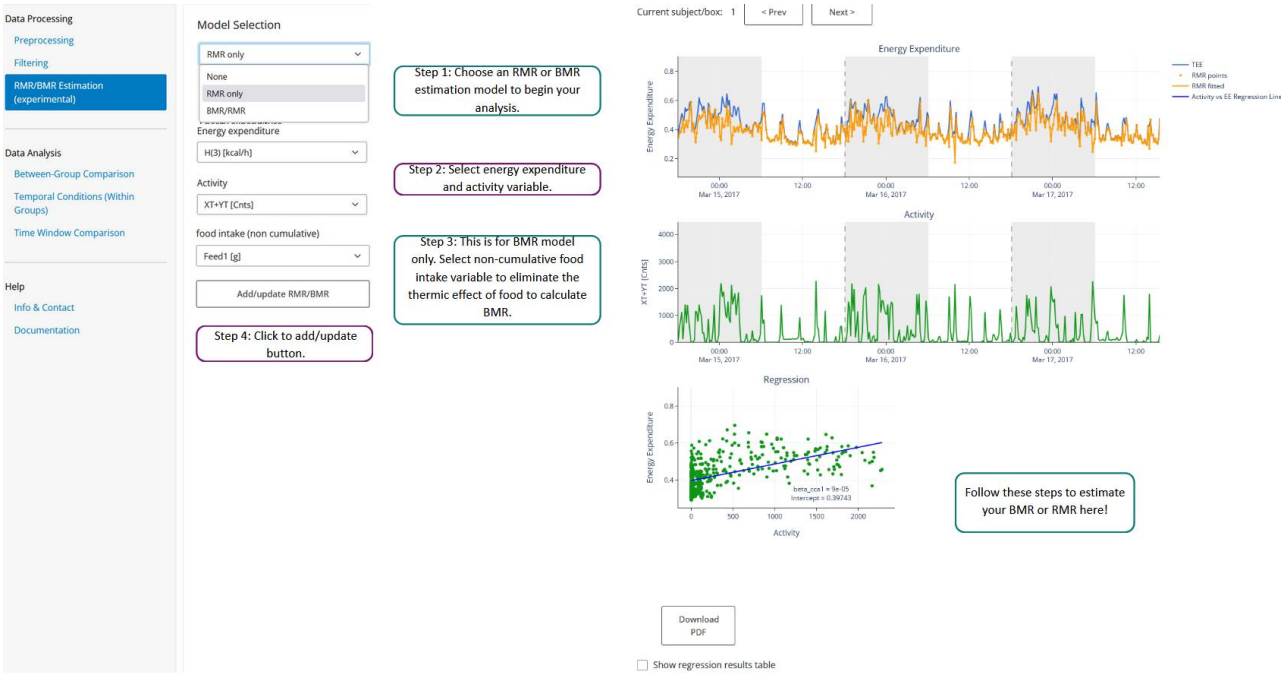


Figure 4: *RMR estimation based on activity data.*

Calopy has a unique feature implemented that allows for the estimation of Resting and Basic metabolic Rates (**BMR/RMR**), based on the work by Klinken et al. [1].

This feature estimates **Resting Metabolic Rate (RMR)** and **Basic Metabolic Rate (BMR)** to analyse energy expenditure. **RMR** represents energy use at rest, while **BMR** is the minimum energy required for essential physiological functions. The estimation is based on linear models that decompose total energy expenditure (**TEE**) into **RMR, BMR, Activity-Related Energy Expenditure (AEE), and the Thermic Effect of Food (TEF)**. A full explanation of the mathematical background can be found at Klinken et al [1], or at the **Calopy** documentation.

Model Selection Guide

1. **RMR Model (Figure 4)**

   This model estimates energy expenditure by separating resting and activity-related energy.

   Steps:

   - Select your energy expenditure variable (e.g., kcal/h).
   - Select your activity variable.
   - Click Add/Update to calculate.

   Outputs:

   - Adds rmr_estimate (time-resolved) to metabolic variables.
   - Adds beta_cca1 and RMR_intercept to metadata.

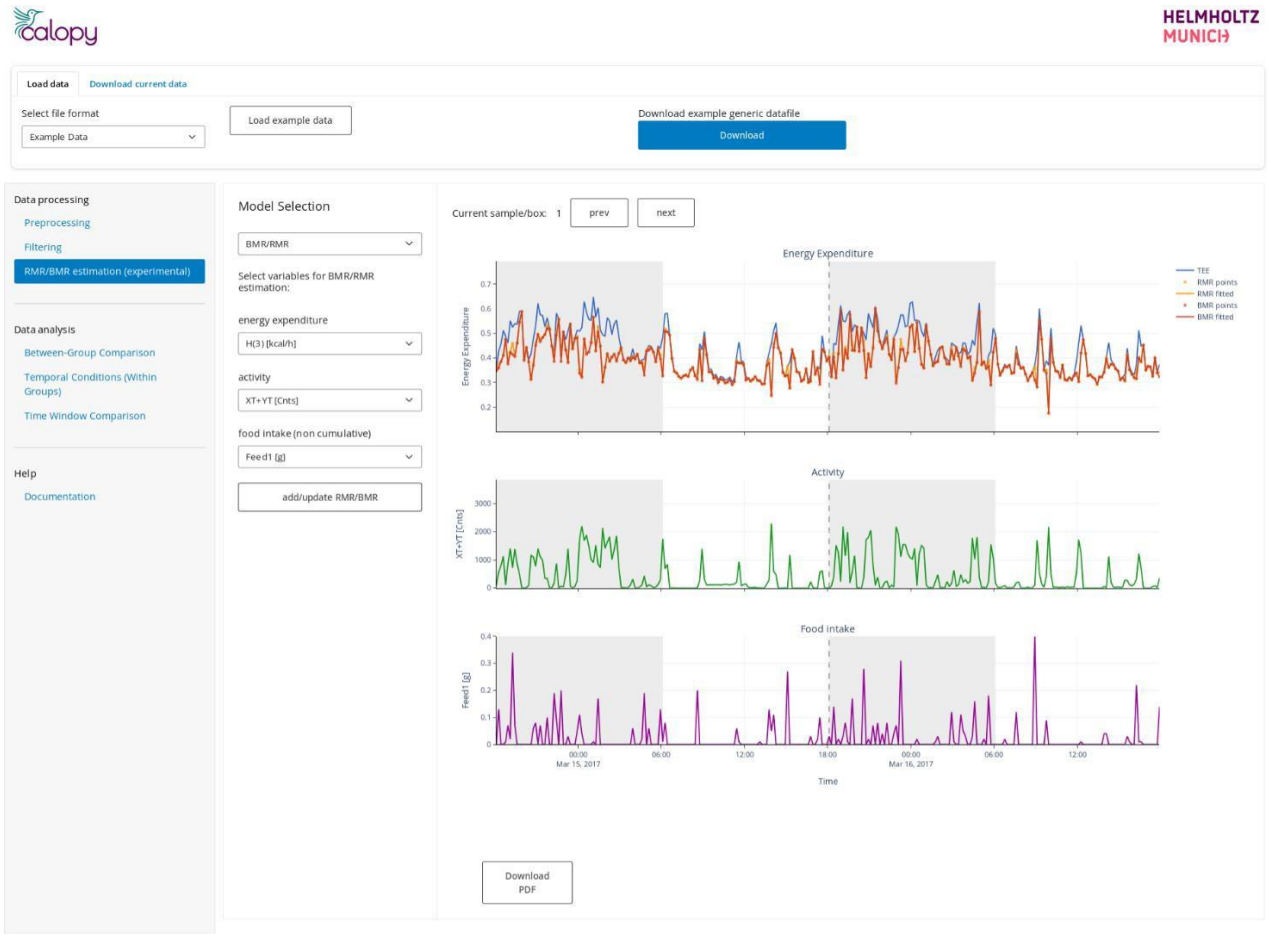


Figure 5: *RMR/BMR estimation based on activity and food intake.*

2. **BMR/RMR Model (Figure 5)**

This model further separates out the thermic effect of food (TEF) using food intake data.

Steps:

- Select your energy expenditure variable (e.g., kcal/h).
- Select your activity variable (dimensionless).
- Select non-cumulative food intake variable.
- Click Add/Update to calculate.

Outputs:

- Adds rmr_estimate and bmr_estimate (time-resolved) to metabolic variables.
- Adds beta_cca2, beta_tef, RMR_intercept, and BMR_intercept to metadata.

These models help break down total energy expenditure into AEE (activity energy expenditure), RMR, BMR, and TEF.

Note: This method has **limitations**, such as reliance on **linear regression** and potential inaccuracies due to incomplete activity or food intake data. Use with care!

DATA ANALYSIS (with use cases)

Currently **Calopy** enables **exploratory analysis of Indirect Calorimetry (IC) data** with three key methods:

1. **Between-Group Comparison** – Analyzes differences in metabolic variables across groups using **one-way or two-way ANOVA, ANCOVA (with covariates like body weight), and linear regression**.
2. **Temporal Conditions (Within Groups)** – Examines metabolic changes within a group over different time conditions, allowing for **repeated measures ANOVA**.
3. **Time Window Analysis** – Segments data into **consecutive or overlapping time windows**, performing **ANOVA-based comparisons** between groups.

Calopy provides **customizable window sizes, flexible variable selection, and day/night phase analysis** to enhance metabolic data interpretation.

Since **Calopy** is designed for **exploratory and flexible data analysis**, we will not offer a step-by-step guide for every scenario but instead provides **key examples of common use cases** to guide your analysis.

Between Group Comparison

The **between-group analysis** tests various dependent variables across different conditions or treatments (i.e., conditional variables). **Calopy's** key strength lies in its flexibility to select and combine predictive and dependent variables:

Dependent Variables

Select metabolic variable
Select the Variable you want to test, you can choose from metabolic variables and continuous metadata.

Filter light/dark phase
Filter for values (and features) only of one phase, either light or dark. Only works with metabolic variables, not with metadata.

Select feature
Select a feature from your metabolic variable. These features strongly depend on the filtering applied during **Data Processing.** Use with care!

Predictive Variable

Select grouping/aggregation
Select the predictive grouping variable.
Note: For regression-based analysis like ANCOVA you must select **Use linear model** from the Options, See below.

Additional Options for Between Group Comparison

In addition to the upper example **Calopy** allows for further options when comparing group effects to continuous variables**:**

TWO-WAY ANOVA
In case your experiment contains two grouping variables e.g. Genotype & Diet you may want to perform a two-way ANOVA. Just select the **Use two-way ANOVA tick-box** and select your additional grouping variable.
Note:

- **TWO-WAY ANOVA is NOT** recommended for testing differences in EE since the use of two-categorical predictive variables and ANCOVA is currently not supported in **Calopy**
- TWO-WAY ANOVA and welch ANOVA cannot be combined

Compare light/dark
To test differences of your dependent variable between light and dark phase select **Compare light/dark tick-box.** This performs a 2way ANOVA including light/dark phases as additional predictive variable/factor. Can only be performed on Metabolic variables.

Note: To compare grouping effects **only for dark or light phases**, use the dark/light dropdown filter instead.

Use Linear models - Linear regression
Beside the previously described feature, **Calopy** offers the unique feature to perform linear regression on all kinds of continuous metadata or features extracted from metabolic variables. Although technically speaking this does not refer to test for between groups comparison, we decided to keep this option here due underline the exploratory character of **Calopy**.

Since possibilities here are manifold, we here describe one exemplary use-case: Predicting Body Weight Gain through metabolic flexibility.

Comparing EE between groups

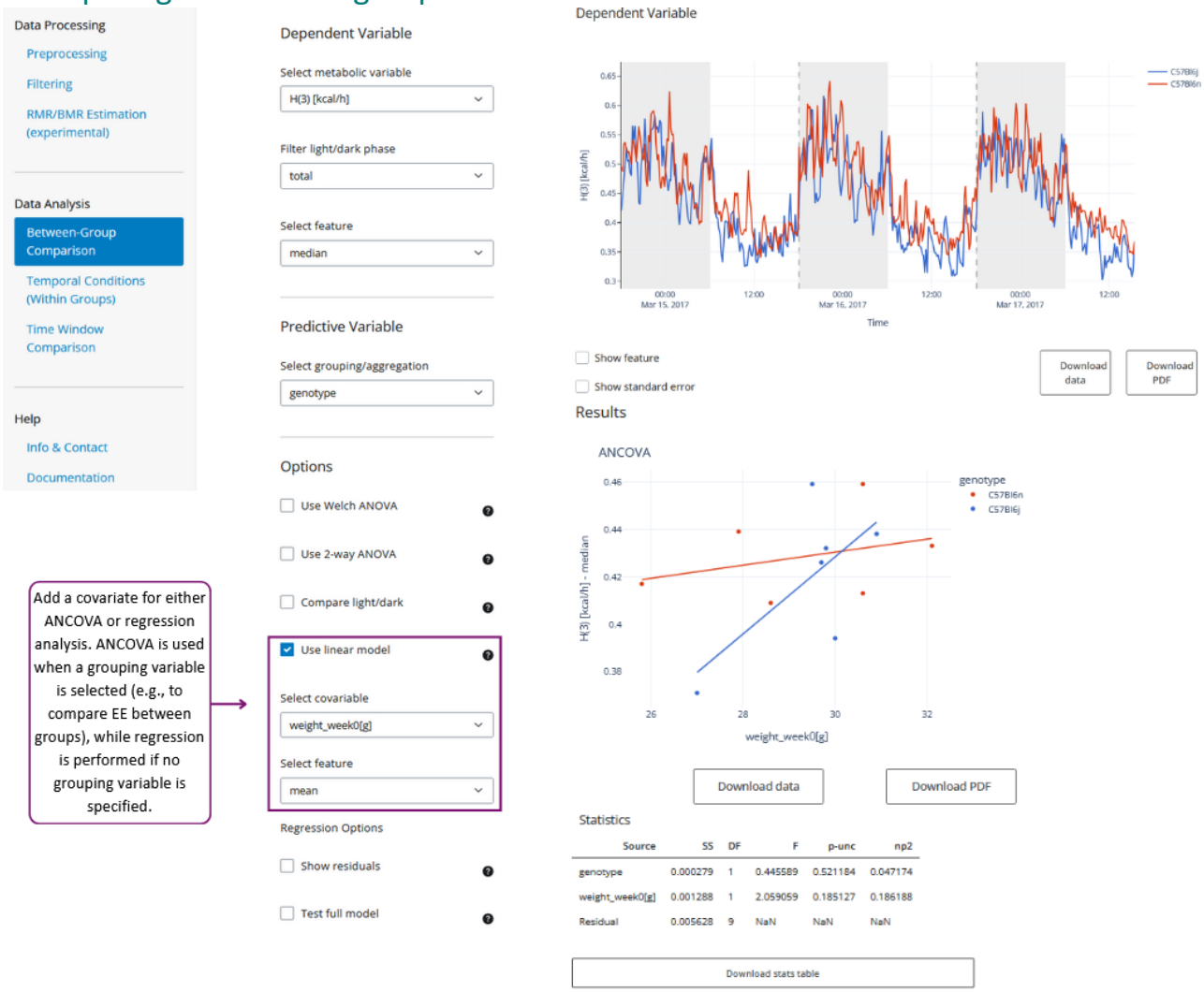


Figure 6: *Comparing EE between groups*

To compare Energy Expenditure (EE) between groups (Mouse strains, diet, drug treatment...) we strongly recommend only using EE or heat production as measured as kcal/h and correct for body weight using ANCOVA.

In **Calopy**, after preprocessing (trimming, eventually sample removing) select **Between-groups comparison**. Select EE variable (here H(3), [kcal/h]) as dependent variable.

You may want to compare EE only for dark/light phase or total. select using light dark filter. As a feature we recommend selecting **mean or median (global)**. Selectin mean or median in 24h split median may be an alternative here.

Select your **grouping variable.** Tick box: **Use linear model** and select **body weight** as covariable. In case your body weight is imported as a time resolved variable you should choose **mean or median** from the feature. In case BW is imported as metadata, the feature selection will have no effect.
Note: Also, **lean mass** may be a good choice as a covariable here [2].

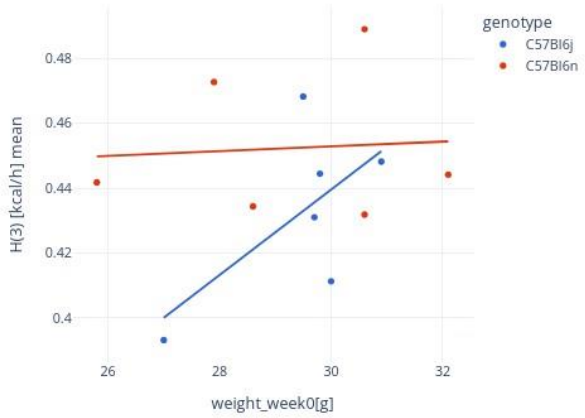
ANCOVA now compares differences in EE between the two groups, by correcting for BW. Thus, your data needs to fulfil several assumptions:

1. The dependent variable and covariate variables are measured on a continuous scale (at the interval or ratio level).
2. The independent variable consists of two or more categorical, independent groups.
3. There is independence of observations (there is no relationship between the observations in each group or between the groups themselves).
4. There are no significant outliers or no points within data that do not follow the regular pattern.
5. The dependent variable is approximately normally distributed within each subpopulation.
6. There is homogeneity of variances (the variance of the dependent variable is equal over all subpopulations).
7. The covariate is linearly related to the dependent variable at each level of the independent variable.
8. There is homogeneity of regression slopes (there is no interaction between the independent variable and the covariate).

Once you have performed your experiment according to standard settings and your system is well calibrated and working correct assumptions 1-6 should be fulfilled.
Also, number 7 holds true to the current assumptions in the field.

Number 8 however is usually violated once the regression lines are not parallel. You test this in **Calopy** be selecting the tick-box ''test full model'' to perform a linear regression including an intersection term (Figure 7). However, although in our model we did not see a significant interaction, we still doubt that the null hypothesis is true due to a lack of statistical power.

Results



Download data

Download pdf

Statistics

Parameter	Coef.	Std.Err.	t	P>|t|	[0.025	0.975]
Intercept	0.044968	0.247730	0.181521	0.860474	-0.526298	0.616235
genotype[T.C57Bl6n]	0.385719	0.285447	1.351281	0.213574	-0.272523	1.043961
weight_week0[g]	0.013152	0.008395	1.566592	0.155844	-0.006208	0.032512
weight_week0[g]:genotype[T.C57Bl6n]	-0.012413	0.009687	-1.281376	0.235953	-0.034752	0.009926

Download stats table

```
Statistics was done using statsmodels[2] (v0.14.4)  'OLS regression' function with 'genotype' and
'weight_week0[g]' as predictive variables including an interaction term.

[1] Vallat, R. (2018). Pingouin: statistics in Python. Journal of Open Source Software, 3(31), 1026,
https://doi.org/10.21105/joss.01026
[2] Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In 9th Python in Sc
https://www.statsmodels.org/
```

Figure 7: *Testing independence for ANCOVA using a full model.*

Comparing continuous Metadata

Figure 8: *Compare Body Weight between groups*

Of course, you can also use **Calopy** to test your continuous metadata using ANOVA (Figure 8). In this example we choose BW as a continuous metadata variable and selected a grouping variable. Most of the other options are without effect here.
Note: You may want to select Welch ANOVA from the options which is the proper choice in case the homogeneity of variances assumption is not met, especially with unequal sample sizes.

Predicting Body Weight Gain through metabolic flexibility:

***Figure 9**: Predicting Body Weight Gain through metabolic flexibility*

Our example data experiment was conducted with 12 mice from two different strains (N=6 per Genotype) which were recorded in our IC system over three days using standard chow diet and subsequently kept over 12 weeks on a high fat died (HFD). Body weight was measured before IC measurement and after 12 weeks HFT.

1. RER was filtered using a "**generalized Additive Model**" to extract denoised features such as minima and maxima.
2. We here selected the weight gain (in %) as the dependent variable (Other settings to depend on variable are ignored here since we use a continuous metadata variable).
3. We use a Linear model to perform regression analysis
4. As predictive variable we choose **RER** and selected **maximum (daysplit/median)** feature.
5. Regression residuals are also shown

Interpretation: From this setting we see a significant predictive potential of the maximal RER -- which can be interpreted as the ability to switch to Carbohydrate energy usage during active phase, thus an indicator for good metabolic flexibility -- to predict weight gain upon a high fat diet. The higher the amplitude, the less the gained weight.

Note: Although **Calopy's** ability to combine filtering and feature extraction with a flexible analysis toolbox enables various comparisons and tests, we strongly recommend that users carefully validate their settings, as improper configurations may lead to misleading or uninterpretable results.

Temporal Conditions

With temporal conditions you can perform tests upon temporal changes during your experiment like changing cage temperature or food.

Dependent Variables

Select metabolic variable
Select the Variable you want to test, you can choose from metabolic variables and continuous metadata.

Select feature
Select a feature from your metabolic variable. These features strongly depend on the filtering applied during **Data Processing.** Use with care!

Conditional Variables

Select grouping/aggregation
Select the predictive grouping variable

Temporal Conditions
Edit Conditions button opens a dialog that allows you to create any number of temporal conditions.

Figure 10: *Comparing VO_2 to different temperature changes over time.*

Our presented example here illustrates how you can compare changes over time by defining temporal conditions. Conditions can be defined and adjusted and adjusted to your experimental design. In this experiment cage temperature was reduced from Thermoneutrality (30°C) to room temperature (23°C) to cold exposure (4°C), each for two days.

Time Window Comparison

Time window comparison is a widely used way to compare your metabolic variables between groups using time windows.

Dependent Variables

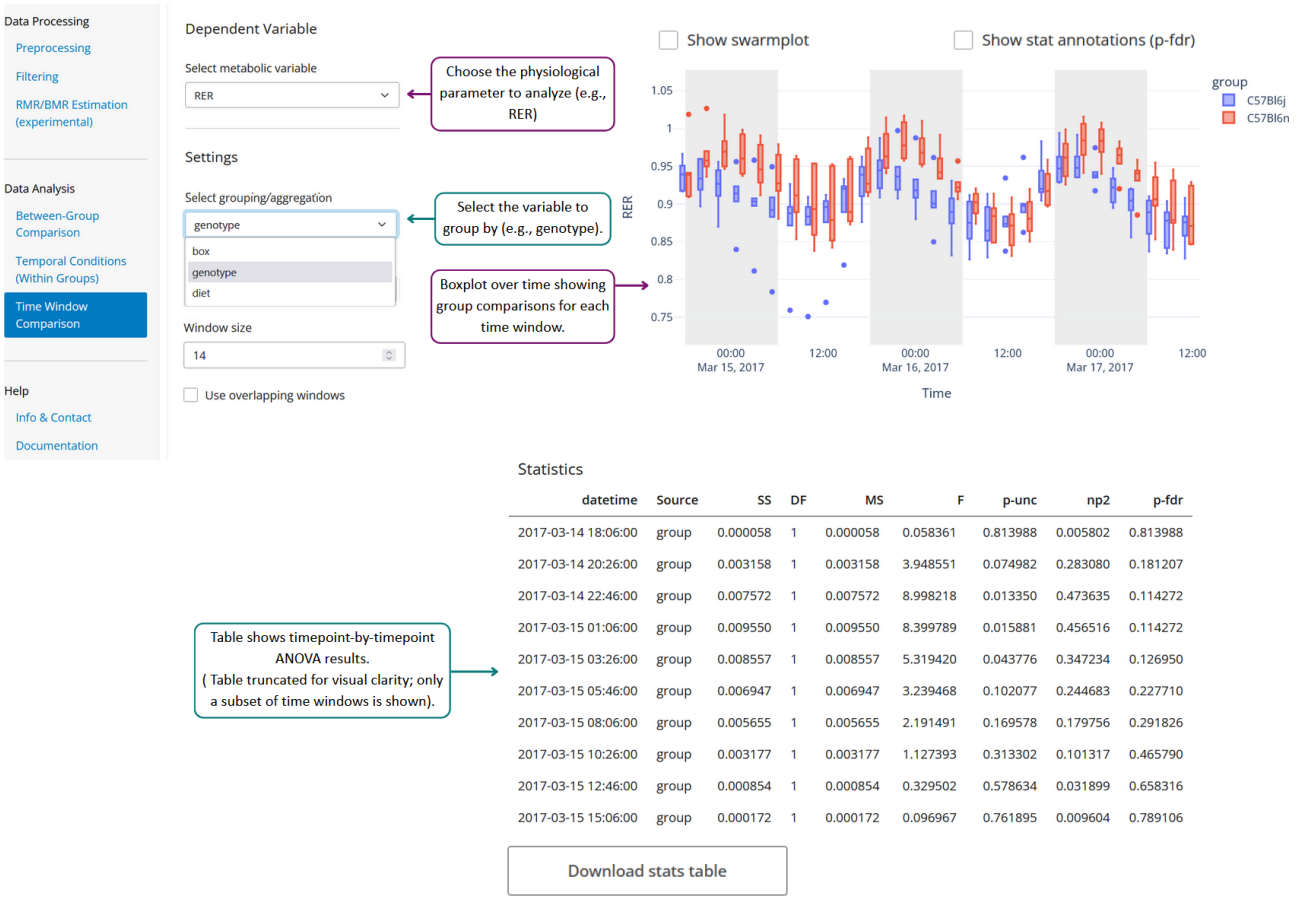
Select metabolic variable
Select the Variable you want to test, you can choose from metabolic variables and continuous metadata.

Settings

Select grouping/aggregation
Select the predictive grouping variable to test the dependent variable for each window.

Window size
Adjust the window size in time sample steps for windows to compare.

Use overlapping windows
Allows for additional settings to adjust time samples between windows (overlapping windows)



datetime	Source	SS	DF	MS	F	p-unc	np2	p-fdr
2017-03-14 18:06:00	group	0.000058	1	0.000058	0.058361	0.813988	0.005802	0.813988
2017-03-14 20:26:00	group	0.003158	1	0.003158	3.948551	0.074982	0.283080	0.181207
2017-03-14 22:46:00	group	0.007572	1	0.007572	8.998218	0.013350	0.473635	0.114272
2017-03-15 01:06:00	group	0.009550	1	0.009550	8.399789	0.015881	0.456516	0.114272
2017-03-15 03:26:00	group	0.008557	1	0.008557	5.319420	0.043776	0.347234	0.126950
2017-03-15 05:46:00	group	0.006947	1	0.006947	3.239468	0.102077	0.244683	0.227710
2017-03-15 08:06:00	group	0.005655	1	0.005655	2.191491	0.169578	0.179756	0.291826
2017-03-15 10:26:00	group	0.003177	1	0.003177	1.127393	0.313302	0.101317	0.465790
2017-03-15 12:46:00	group	0.000854	1	0.000854	0.329502	0.578634	0.031899	0.658316
2017-03-15 15:06:00	group	0.000172	1	0.000172	0.096967	0.761895	0.009604	0.789106

Figure 11: *Time window comparison*

Time Window Analysis lets you select a depended metabolic variable, a grouping factor and a time window size, measured in sample rate. You may want to use overlapping or temporal offset windows which you may set when selecting **use overlapping windows** tick-box.

REFERENCES

[1] Van Klinken JB, et al. Estimation of activity-related energy expenditure and resting metabolic rate in freely moving mice from indirect calorimetry data. *PLoS One*. 2012;7(5):e36162.
DOI: https://doi.org/10.1371/journal.pone.0036162

[2] Virtue, S., Lelliott, C.J. & Vidal-Puig, A. What is the most appropriate covariate in ANCOVA when analysing metabolic rate?. *Nat Metab* **3**, 1585 (2021).
DOI: https://doi.org/10.1038/s42255-021-00505-5