**Introduction to greenfeedr**

**18-21 March, 2025**

**Lecture: 10 AM – 11:30 AM Tuesday**

**Lab: 10 AM – 11:30 AM Friday**

Instructor:

**Guillermo Martinez-Boggio** ([guillermo.martinezboggio@wisc.edu](mailto:guillermo.martinezboggio@wisc.edu))

Required Software: R and RStudio

Course objectives:

* Master the usage of the greenfeedr package.
* Gain hands-on training for efficient GreenFeed data handling.

Expectations of students:

* Install R and RStudio before class time.
* Students should have prior experience with GreenFeed system.

Modes of instruction:

* Virtually classes.

**Program**

The program is designed to provide training on handling and analyzing GreenFeed data using the greenfeedr R-package. It does not cover topics related to GreenFeed theory, operating principles, mechanical design, or maintenance. Participants are encouraged to have prior experience with GreenFeed systems to maximize the benefits of the training.

1. **Downloading data**

The greenfeedr package has two functions, get\_gfdata and report\_gfdata, that allow users to download their processed gas records daily.

**What is the GreenFeed data?** The raw sensor data is processed daily by the C-Lock processor and available online with a one-day delay. C-Lock Inc. provides all users with summarized data for a specific unit/s and time range (upon request in 3-4 weeks). However, using login information, users can retrieve data processed daily through the website (<https://ext.c-lockinc.com/greenfeed.php>) or an application programming interface (API). API is a software interface that allows computers or computer programs to connect.

**What are the differences between daily and summarized data?** Daily data is all gas records processed by the C-Lock processor daily that do not have the extra (manual) check done by the C-Lock team when requested. Summarized data is all gas records processed by the C-Lock processor and with an extra manual check by the C-Lock team provided to users. The differences between data are minimal (one more or fewer records and minor adjustments on gas records) in most cases, and the data table structure is almost the same.

**What is the relevance of daily data?** Users can follow their trials by checking the records generated by GreenFeed. Also, by the end of the trial, users can have some “preliminary” data to explore before getting the summarized data from the C-Lock team. But what is the difference between visits and records? A visit refers to when an animal enters its head into the unit, receives food drops, and exits the unit. A visit record refers to when the visit is translated on a record in the data due to some criteria (e.g., >2 minutes).

We will compare two datasets: one containing daily data and the other consisting of summarized data, using the function compare\_gfdata. C-Lock: Do you want to give users any information regarding data processing once the user requests the final report?

1. **Reporting data**

The greenfeedr package has the function report\_gfdata, which allows users to download and generate a PDF report of the daily or summarized data. Users can also automate the generation of daily PDF reports that include key metrics and plots to check their unit/s.

**What is the relevance of daily reports?** Whilechecking units through the web interface is crucial, understanding how the data is generated is equally important. It is useless if cows visit the unit, but those visits aren’t translated into records. This PDF report offers a valuable summary of the trial's progress and performance. It helps identify days when the GreenFeed system encountered issues and serves as a reference for addressing future questions, such as why certain days had limited cow activity or other anomalies. The report ensures better accountability and traceability throughout the trial by providing this level of detail.

**How to use the report\_gfdata function?** The greenfeedr package has the report\_gfdata function, which allows users to download and generate a report of their processed gas records. What options are available? How do I create a list of trials and generate all reports?

We will work on creating and reading the reports. We will also provide a script to generate many reports in one run. Windows users can have issues with PDFlatex!

1. **Processing data**

A crucial step for analysis is to do correct data processing. We developed the function process\_gfdata to help users process their daily and summarized GreenFeed data.

**What records are removed from the data?** The function process\_gfdata removes all ‘Unknown ID’, records with a low airflow (<25 l/s), and outlier records define as values +- 3SD. However, users can change the default value using the parameter ‘cutoff’. What is the difference between using 1.5\*IQR or SD? We assume the gas records have a normal distribution, plus the checks done by C-Lock (quality control?).

**What is the filtering process?** The GreenFeed systemrelies on animals having multiple records across multiple days. So, it is crucial to check the reliability of our data. We defined three parameters (param1, param2, and min\_time) that allow users to control the data processing.

* Parameter 1 (param1) is the number of records per day.
* Parameter 2 (param2) is the number of days with records in a week.
* Parameter 3 (min\_time) is the minimum duration of an animal’s visit record.

**What is the best combination of parameters?** The best combination depends on the specific characteristics of your trial, such as the season, training protocol, GreenFeed setup, and other contextual factors. There is no universal solution that works for all cases. Using the process\_gfdata function, users can explore various parameter combinations and make informed decisions based on the unique aspects of their data.

We will show an example of how to evaluate the impact of using different ‘threshold’ parameters by looping over a parameter space with 60 or more combinations (param1: 2 to 4, param2: 3 to 7, and min\_time: 2 to 5).

**How are the averages computed?** Daily and weekly weighted averages using visit time. We will discuss the implications of computing a weighted mean for daily and weekly averages instead of the mean. Why give more “value” to the long gas measurements? Does this have any impact on the resulting average?