

# Building a Global Integrated Methane Data Ecosystem

Harmonizing Measurement, MIR, and Genomic Data to Accelerate Livestock Methane Mitigation

**Hanlu Zhang**, Puchun Niu, Miel Hostens

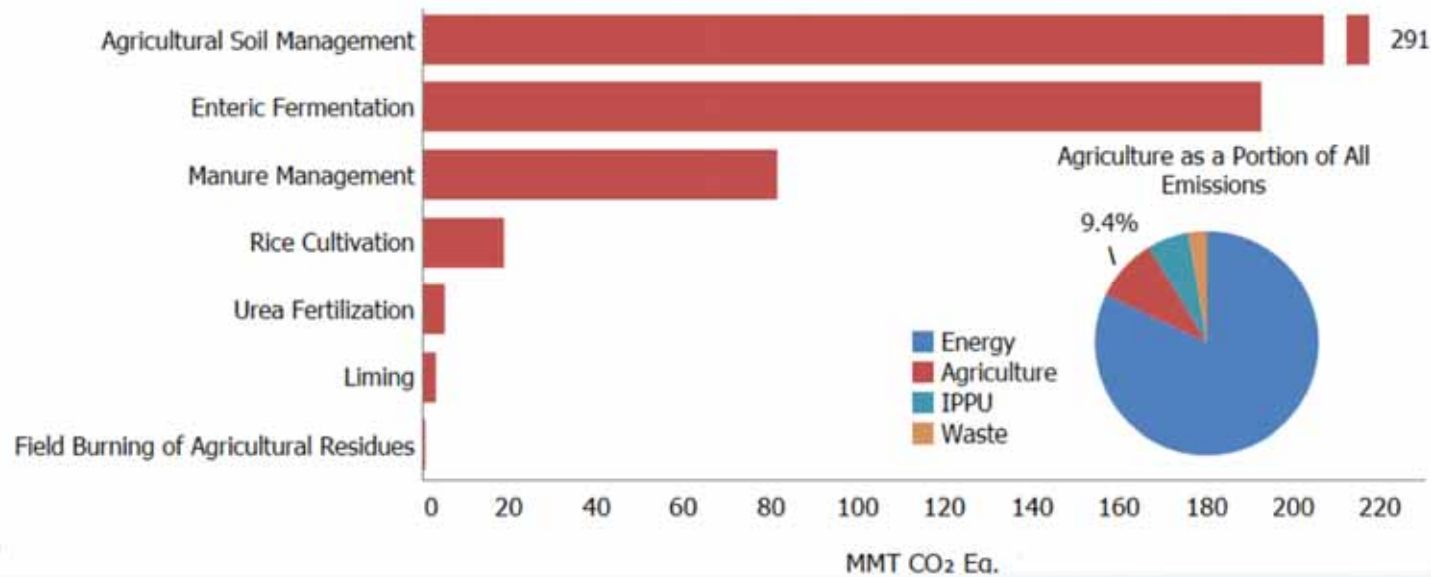
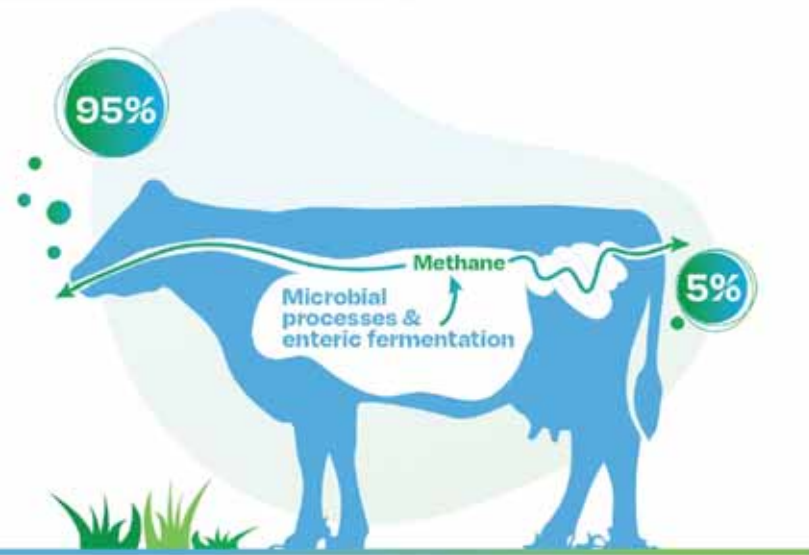
Cornell University

International Committee for Animal Recording (ICAR) • Verona, Italy

# Background

<https://www.dairyconservation.org/enteric-emissions>

## The Enteric Methane Challenge



United States Environmental Protection Agency. 2024. Draft Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2022. <https://www.epa.gov/ghgemissions/draft-inventory-us-greenhouse-gas-emissions-and-sinks-1990-2022>.

**Enteric methane is the single largest source of GHG emissions in the fluid milk chain (USDA, 2021).**

## Background

# Enteric Methane Mitigation



Accurate  
measurement



Moving from estimates  
to high-fidelity, real-  
time data



Standardizing protocols  
Establishing baselines  
Reducing uncertainty



Scalable  
prediction tools



Lowering the cost of  
monitoring by using  
proxies



Developing models  
Farm-level reporting  
Management systems



Coordinated global  
collaboration



Harmonizing  
data  
across borders



Model accuracy  
Aligning industry  
Scaling projects

## Background



Accurate  
measurement



Scalable  
prediction tools



Coordinated global  
collaboration

## The Challenge

### ⚠️ **Fragmented Research**

- ❌ **Siloed Pipelines:** Fragmented across sensors, biology, and institutes.
- ❌ **Data Duplication:** Inefficient use of limited global research funds.
- ❌ **Isolated Assets:** Missing alignment between gas data and multi-omics.

### ➡️ **Mitigation Objectives**

- ✅ **Statistical Power:** Essential to identify low-methane genetics.
- ✅ **Unified Metrics:** Comparable baseline phenotypes across countries.
- ✅ **Scalable Proxies:** Deployable, low-cost proxies on commercial herds.

# Core Research Objectives



## 1. Harmonize Phenotypes

- ✓ Standardize measurement
- ✓ Establish procedural benchmarks
- ✓ Improve global comparability



## 2. Integrate Multi-Modal data

- ✓ Mid-infrared (MIR) milk spectra
- ✓ High-density animal genotypes
- ✓ Dietary & performance covariates

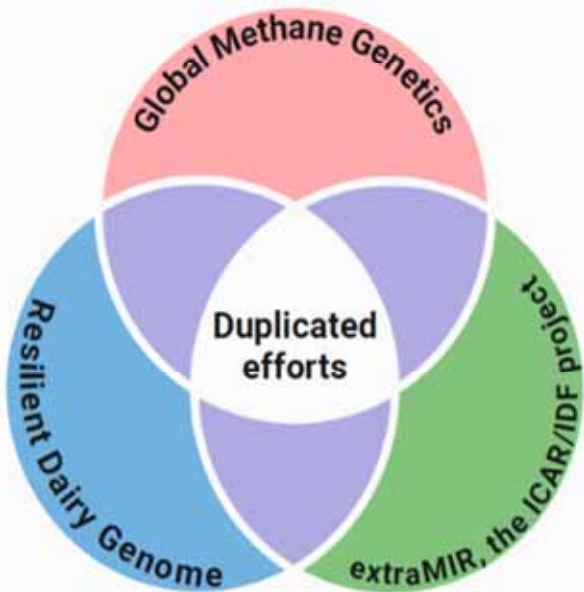


## 3. Predictive Pipelines

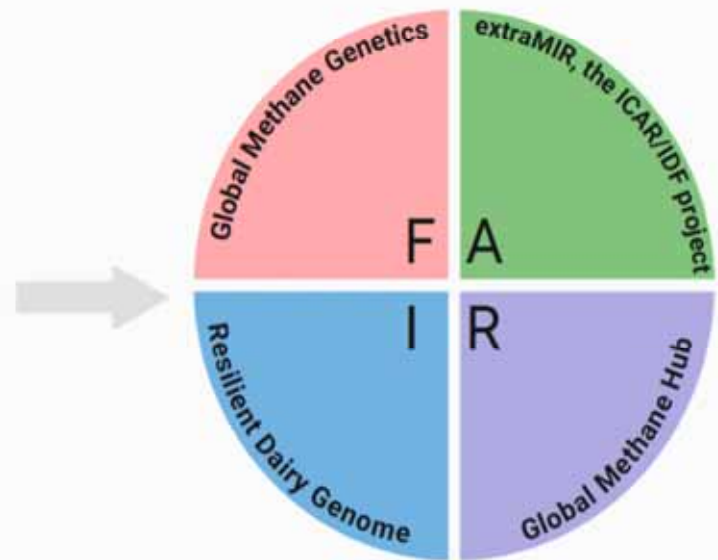
- ✓ Shared cloud workflows
- ✓ Robust prediction models
- ✓ Low-methane genetic selection

# Objectives

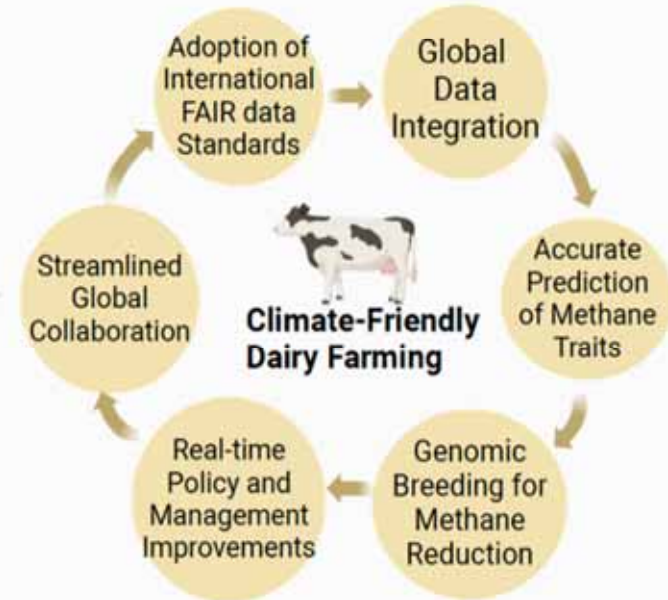
GLOBAL METHANE INITIATIVES



COORDINATED METHANE INITIATIVES



STRATEGIC METHANE MITIGATION



## Objectives



Accurate measurement



Scalable prediction tools



Coordinated global collaboration

### Primary institutions



### Additional Contributors



## The GEMS Foundation

### Scaling Up Methane Recording



Our ongoing initiative expands **GEMS** (*Accurate Gas Emissions Measures from Cattle with the GreenFeed System*).

We are establishing an international data consortium that bridges raw methane gas phenotypes with MIR, genomic and management profiles.



## Current stage

# Global Traction & Engagement

# 30+

Global Cooperating  
Organizations

## Active Global Support

Within the first 2 months: **13 major organizations** signed

Data from **~80 distinct animal research studies** is actively flowing into our pipeline.

**GreenFeed Emissions Measurement System**

A global collaboration using the GreenFeed system to track methane and other gas emissions from ruminants, building a standardized, FAIR data warehouse for climate-smart livestock research.

**OUR MISSION**  
Develop science-based, standardized operating procedures for both utilizing and interpreting GreenFeed data under different management practices - so every partner can produce comparable, defensible emissions measurements.

Global Methane Hub | Cornell University | lead coordinator  
50+ partner institutions | Real-time emissions data

### Consortium Members

Cornell University | University of California | University of Guelph  
ETH Zurich | University of New England | Agriculture and Agri-Food Canada





**10** Partner institutions | **28** Gold studies | **1,564** Animals | **Nov 2014 - Jan 2026** Date range

View partner institutions ▾ | View gold study institutions ▾







## Objectives

# FAIR Data Infrastructure

-  **Centralized Data Lake**  
Aggregates raw gas emissions, feed intakes, MIR spectrums...
-  **Legal Frame (DUA)**  
Standardized Data Usage Agreements protect institutional IP while easing sharing.
-  **Automated Validation**  
Programmatic quality checks to sanitize inputs, adjust sensor drift, and flag anomalies.
-  **ICAR-Aligned Interoperability**  
Maps disparate international databases into one consistent, standardized recording matrix.

## Current stage

# Integrated Data Matrix

Data Layer	Measurement Methods	Primary Purpose	Harmonization Status
Methane Phenotypes	GreenFeed, Respiration Chambers, SF6	Direct gas emission quantification	 Active Collection
Milk Spectra	Mid-Infrared (MIR) Spectroscopy	Scalable proxy-based predictions	 Pipeline Integrated
Genomics	High-Density Genotyping Platforms	Animal breeding & genetic evaluations	 DUA Developing
Metadata	Diet Details, Intake & Production	Covariate adjustment & nutritional modeling	 Active Collection

## Final outcome

# Integrated Biological Layers



### 1. Mid-Infrared (MIR) Spectra

Captured standardly in milk recordings to build low-cost, scalable proxies for individual methane output.



### 2. Genotypes & Selection

Connects physical emissions to genotypes, increasing the accuracy of genomic breeding indexes.



### 3. Cohort Metadata

Metadata are mapped, eliminating database biases during prediction modeling.

$$\text{CH}_4 = f(\text{MIR Spectroscopy} + \text{Genomics (GEBV)} + \text{Nutritional Covariates})$$



## Q & A

Questions, suggestions, willing to collaborate?

 [\*\*\*gems@cornell.edu\*\*\*](mailto:gems@cornell.edu)

Join us in building a unified global methane data ecosystem to  
accelerate livestock mitigation.