Creation of database for meta-analysis
Researchers Sharing Data Was Supposed to Change Science Forever. Did It?

By Lily Hay Newman

The reality of open-source data is a jumbled mess.
Databases

• Why do we need big data?
  – “We have tiny little brains. We can’t understand the big stuff anymore” – The Defense Advanced Research Projects Agency

• The hard thing is not actually to dump your data into the public domain. It’s to dump it in an intelligible way. To make data from a project usable, it takes about 20 percent of a researcher's total work.
Databases

• Creation of the database will be the first step towards analysis
• Once the contracts are signed all participants will be asked to provide their data
• The data will be held under strict confidence and access will be discussed and finalized today
Databases

Announcements

There are currently no active announcements. To add a new announcement, click "Add new announcement" below.

Latest Document Updates

- [OneNote DeletedPages. one](#) - 07/16/2014 11:07 AM
- [Standups. one](#) - 07/16/2014 11:07 AM
- [My Account Rev. pdf](#) - 07/15/2014 12:19 PM
- [Allen’s summary of relationship between decision support & reporting](#) - 07/15/2014 08:26 AM
- [Farm Smart Module Documentation](#) - 07/14/2014 09:18 AM

Global GHG Network

Dairy Research Institute

Sustainability
Dairy Research Institute
Environmental
Dairy Research Institute
Discussions
Lists
Libraries
Calendar
Documents
Farm Smart Team
Feed Loss
Energy
Nutrient Analysis
Herd
LCA 1.5 Refactor
Data Dictionary
Best Management Resources

CheeseTesting

MODELING SUSTAINABLE AGRICULTURE at UC DAVIS
Databases - Australia

NLMP National Livestock Methane Database

- Projects
  - What are Projects?
  - Projects are used to create, manage and publish collections of experiments and associated datasets.
  - Projects can be managed by a number of organisations jointly, but there is normally a single organisation that leads the research work.
  - The project category allows a user to download a copy of the project final report.

- Add Project
  - Search projects:

- 72 projects found
  - Order by: Name Ascending

- RELRP: A genomic strategy to identify archaeal viruses in the rumen
  - Phage therapy is becoming increasingly important as a means of eradicating or...
  - 1 Experiment

- RELRP: Application and extension of FarmGAS decision support tool – trainer to trainer program
  - The free online FarmGAS Scenario Tool was developed with funding from the...
  - 1 Experiment

- RELRP: Blood

- RELRP: Antimethanogenic bioactivity of Australian plants for grazing systems
  - In vitro information has been collected on bioactive properties of over 130...
  - 1 Experiment

- RELRP: Archaeaphage therapy to control rumen methanogens
  - Phage therapy is becoming increasingly important as a means of eradicating or...
  - 1 Experiment

- RELRP: Breeding for...
Databases - Australia

![NLMP Database Interface]

The NLMP (National Landfill Methane Program) Database provides information on landfill methane emissions and can be accessed through various tools and datasets.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Site</th>
<th>Collection</th>
<th>Diet</th>
<th>Animal No.</th>
<th>AccOH</th>
<th>PV</th>
<th>Score</th>
<th>d14C</th>
<th>d13C</th>
<th>d18O</th>
<th>VelMax</th>
<th>hVel</th>
<th>MaxFlow</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLMP 1</td>
<td>Belmont</td>
<td>1</td>
<td>April 2013</td>
<td>120412</td>
<td>33.73</td>
<td>7.99</td>
<td>0.92</td>
<td>5.76</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLMP 2</td>
<td>Belmont</td>
<td>1</td>
<td>April 2013</td>
<td>120485</td>
<td>61.56</td>
<td>11.51</td>
<td>1.80</td>
<td>7.71</td>
<td>0.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLMP 3</td>
<td>Belmont</td>
<td>2</td>
<td>April 2013</td>
<td>120496</td>
<td>49.37</td>
<td>9.65</td>
<td>1.92</td>
<td>6.59</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLMP 4</td>
<td>Belmont</td>
<td>3</td>
<td>April 2013</td>
<td>120506</td>
<td>51.31</td>
<td>16.31</td>
<td>0.75</td>
<td>4.78</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLMP 5</td>
<td>Belmont</td>
<td>3</td>
<td>April 2013</td>
<td>120516</td>
<td>42.22</td>
<td>8.41</td>
<td>0.58</td>
<td>5.19</td>
<td>0.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NLMP 6</td>
<td>Belmont</td>
<td>3</td>
<td>April 2013</td>
<td>120521</td>
<td>39.31</td>
<td>7.52</td>
<td>0.92</td>
<td>7.52</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This dataset can be accessed through the NLMP website, providing valuable insights into landfill methane emissions.
Databases – Treatment Means

• At the minimum we need information on:
  – Number of observations
  – Some measure of variability. SD is preferred but SEM or SED will also be usable by converting it to SD
  – Measurement methodology (e.g., VHC, GreenFeed etc for methane)

• Need to conduct ‘quality control’ to make sure numbers are within the expected range
Analysis

• A meta-analytical approach will be used.
• All analysis will be conducted using R statistical software or WinBugs
• A correlation matrix will be developed in order to avoid multi-collinearity issues
• Variable selection will be conducted using reversible jump Markov Chain Monte Carlo method
Bayesian Hierarchical Model

- Model for the data given model parameters
  - Let $y_{ijk}$ denotes the $k^{th}$ ($i = 1, \ldots, n_{ij}$) record on the $i^{th}$ ($i = 1, \ldots, I$) animal in the $j^{th}$ ($j = 1, \ldots, J$) study

$$y_{ijk} | \beta, \alpha_i, \delta_j, \tau, \nu \sim t(X'_{ijk} \beta + \alpha_i + \delta_j, \tau, \nu)$$

- Data modeled through a student’s-$t$ density
- Expected value modeled through covariates selected by RJMCMC plus animal and study random effects
## Selected Models

### MCMC Posterior means

<table>
<thead>
<tr>
<th>Model</th>
<th>Posterior Prob.</th>
<th>Prediction Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>GE</td>
<td>-</td>
<td>$\text{CH}_4 = 3.25 (0.429) + 0.043 (0.001) \times \text{GEI}$</td>
</tr>
<tr>
<td>Dietary</td>
<td>0.74</td>
<td>$\text{CH}_4 = 0.225 (0.7133) + 0.042 (0.001) \times \text{GEI} + 0.125 (0.015) \times \text{NDF} - 0.329 (0.094) \times \text{EE}$</td>
</tr>
<tr>
<td>Animal</td>
<td>0.86</td>
<td>$\text{CH}_4 = -9.31 (1.06) + 0.042 (0.001) \times \text{GEI} + 0.094 (0.014) \times \text{NDF} - 0.381 (0.092) \times \text{EE} + 0.0078 (0.001) \times \text{BW} + 1.62 (0.119) \times \text{MF}$</td>
</tr>
</tbody>
</table>

Moraes et al. 2014, Global Change Biology