Review: To be or not to be an identifiable model. Is this a relevant question in animal science modelling?

R. Muñoz-Tamayo, L. Puillet, J.B. Daniel, D. Sauvant, O. Martin, M. Taghipoor, and P. Blavy

**Supplementary material S2**

**Structural identifiability of a ruminal lipolysis and biohydrogenation model**

To enlarge our discussion on structural identifiability, we analyse here two mathematical models developed by [Moate *et al.*, 2008](#_ENREF_7) to represent *in vitro* kinetics of two biological processes namely ruminal lipolysis and biohydrogentation. These two multistep biological processes were described mathematically by using a multi-compartmental modelling approach. Each model was calibrated by published experimental data ([Noble *et al.*, 1974](#_ENREF_9)) where only the biological process of interest took place. Reactions rates were defined by either Michaelis-Menten kinetics or first-order kinetics.

For the lipolysis model, the set of ordinary differential equations of the model are

$$\frac{dx\_{1}\left(t\right)}{dt}=-\frac{k\_{1}∙x\_{1}}{k\_{2}+x\_{1}}∙e^{-k\_{3}∙t}, x\_{1}\left(0\right)=x\_{10} $$

$$\frac{dx\_{2}\left(t\right)}{dt}=\frac{2}{3}∙\frac{k\_{1}∙x\_{1}}{k\_{2}+x\_{1}}∙e^{-k\_{3}∙t}-k\_{4}∙x\_{2}, x\_{2}\left(0\right)=0 $$

$$\frac{dx\_{3}\left(t\right)}{dt}=\frac{1}{2}∙k\_{4}∙x\_{2}-k\_{4}∙x\_{3}, x\_{3}\left(0\right)=0 $$

$\frac{dx\_{4}\left(t\right)}{dt}=\frac{1}{3}∙\frac{k\_{1}∙x\_{1}}{k\_{2}+x\_{1}}∙e^{-k\_{3}∙t}+\frac{1}{2}∙k\_{4}∙x\_{2}+k\_{4}∙x\_{3}, x\_{4}\left(0\right)= 0$ (1)

where $x\_{1}$ is the concentration of triglyceride fatty acids, $x\_{2}$ is the concentration of diglyceride fatty acids , $x\_{3}$ is the concentration of monoglyceride fatty acids and $x\_{4}$ is the concentration of non-esterified fatty acids. From the experimental setup, the model observables are

$$y\_{1}(t)= x\_{1}(t)$$

$$y\_{2}\left(t\right)= x\_{2}\left(t\right)+x\_{3}\left(t\right)$$

 $y\_{3}(t)= x\_{4}(t)$ (2)

The observable $y\_{2}\left(t\right)$ is an aggregated pool of mono and diglyceride fatty acids

named by [Noble *et al.*, 1974](#_ENREF_9) as partial glycerides.

Identifiability analysis was performed using DAISY ([Bellu *et al.*, 2007](#_ENREF_3)), GenSSI ([Chis *et al.*, 2011](#_ENREF_4)) and IdentifiabilityAnalysis ([Karlsson *et al.*, 2012](#_ENREF_5)). The model parameters are identifiable. For all of these software tools, the computation time was less than one second on an Intel processor of 3.20 GHz with 8.0 GB RAM.

Interestingly, we determined that the observable $y\_{2}\left(t\right)= x\_{2}\left(t\right)+x\_{3}\left(t\right)$ was not necessary to guarantee the identifiability of the model parameters. This result indicates that sometimes having many observations does not imply necessarily an improvement on the structural identifiability of a model. Indeed, in a context of resource-consuming measurements, we might be interested in identifying a minimal set of measurements that guarantee structural identifiability ([Anguelova *et al.*, 2012](#_ENREF_1)). However, it is clear that in practice having more measurements can be instrumental for performing model calibration.

For the biohydrogenation model, the set of ordinary differential equations of the model are

$$\frac{dx\_{4}\left(t\right)}{dt}=-\frac{k\_{5}∙x\_{4}}{k\_{6}+x\_{4}}, x\_{4}\left(0\right)=x\_{40} $$

$$\frac{dx\_{5}\left(t\right)}{dt}=\frac{k\_{5}∙x\_{4}}{k\_{6}+x\_{4}}-\frac{k\_{7}∙x\_{5}}{k\_{8}+x\_{5}+x\_{6}}, x\_{5}\left(0\right)=0 $$

$$\frac{dx\_{6}\left(t\right)}{dt}=\frac{k\_{7}∙x\_{5}}{k\_{8}+x\_{5}+x\_{6}}-k\_{9}∙x\_{6}\frac{k\_{10}-x\_{6}}{k\_{10}}, x\_{6}\left(0\right)= 0 $$

$\frac{dx\_{7}\left(t\right)}{dt}=k\_{9}∙x\_{6}\frac{k\_{10}-x\_{6}}{k\_{10}}, x\_{7}\left(0\right)= 0 $(3)

where $x\_{4}$ is the concentration of non-esterified fatty acids (linoleic), $x\_{5}$ is the concentration of rumenic acid, $x\_{6}$ is the concentration of vaccenic and $x\_{7}$ is the concentration of stearic acid. From the experimental setup, the model observables are

$$y\_{1}(t)= x\_{4}(t)$$

$$y\_{2}\left(t\right)= x\_{5}\left(t\right)$$

$$y\_{3}\left(t\right)= x\_{6}\left(t\right)$$

 $y\_{4}(t)= x\_{7}\left(t\right)$ (4)

Structural identifiability analysis was performed with DAISY. The model parameters are identifiable. The computation time was less than one second.

It should be noted however, that the accuracy of the parameter estimates strongly depends on the quality of the available data for calibration. Indeed, [Moate *et al.*, 2008](#_ENREF_7) encountered practical identifiability problems for estimating some model parameters. To circumvent this obstacle, some parameters were fixed and set as known values.

Note that a mathematical model representing both lipolysis and biohydrogenation should integrate the model equations in Eq. (1) and Eq. (3). For the non-esterified fatty acids ($x\_{4}$), the resulting differential equation is

$\frac{dx\_{4}\left(t\right)}{dt}=\frac{1}{3}∙\frac{k\_{1}∙x\_{1}}{k\_{2}+x\_{1}}∙e^{-k\_{3}∙t}+\frac{1}{2}∙k\_{4}∙x\_{2}+k\_{4}∙x\_{3}-\frac{k\_{5}∙x\_{4}}{k\_{6}+x\_{4}}, x\_{4}\left(0\right)= x\_{40} $(5)

The full model integrating both lipolysis and biohydrogenation has seven state variables and ten parameters. As an academic exercise, we analysed the structural identifiability of the full model using the model observables in Eq. (2) and Eq. (4). The model parameters were identifiable. The computation time in DAISY was about 30 seconds, indicating the computational effort required when model complexity increases (30 seconds vs 1 second).

The results from structural identifiability analysis are encouraging for modelling attempts towards an enhanced mechanistic representation of the rumen ecosystem. Although improvements are needed (in particular for describing biohydrogenation), the model developed by ([Moate *et al.*, 2008](#_ENREF_7)) provided a parsimonious and biological based approach that can be used as scaffold for incorporating lipid metabolism in existing models of rumen fermentation ([Baldwin *et al.*, 1987](#_ENREF_2); [Mills *et al.*, 2001](#_ENREF_6); [Muñoz-Tamayo *et al.*, 2016](#_ENREF_8)). In a scenario of constructing a predictive model of rumen fermentation, the question of identifiability of the model of lipolysis and biohydrogenation is relevant since parameter estimates obtained from *in vitro* data can be used as priors in an extended model describing the *in vivo* system.

References

Anguelova M, Karlsson J and Jirstrand M 2012. Minimal output sets for identifiability. Mathematical Biosciences 239, 139-153.

Baldwin RL, Thornley JH and Beever DE 1987. Metabolism of the lactating cow. II. Digestive elements of a mechanistic model. J. Dairy Res. 54, 107-131.

Bellu G, Saccomani MP, Audoly S and D'Angio L 2007. DAISY: A new software tool to test global identifiability of biological and physiological systems. Computer Methods and Programs in Biomedicine 88, 52-61.

Chis O, Banga JR and Balsa-Canto E 2011. GenSSI: a software toolbox for structural identifiability analysis of biological models. Bioinformatics 27, 2610-2611.

Karlsson J, Anguelova M and Jirstrand M 2012. An efficient method for structural identifiability analysis of large dynamic systems. In 16th IFAC Symposium on System Identification pp. 941-946.

Mills JA, Dijkstra J, Bannink A, Cammell SB, Kebreab E and France J 2001. A mechanistic model of whole-tract digestion and methanogenesis in the lactating dairy cow: model development, evaluation, and application. J. Anim. Sci. 79, 1584-1597.

Moate PJ, Boston RC, Jenkins TC and Lean IJ 2008. Kinetics of ruminal lipolysis of triacylglycerol and biohydrogenation of long-chain fatty acids: New insights from old data. Journal of Dairy Science 91, 731-742.

Muñoz-Tamayo R, Giger-Reverdin S and Sauvant D 2016. Mechanistic modelling of in vitro fermentation by rumen microbiota. Anim. Feed Sci. Technol. 220, 1-21.

Noble RC, Moore JH and Harfoot CG 1974. Observations on the pattern on biohydrogenation of esterified and unesterified linoleic acid in the rumen. Br J Nutr 31, 99-108.