



HAL
open science

Use and misuse of meta-analysis in Animal Science

Daniel Sauvant, M P Letourneau-Montminy, Philippe Schmidely, Maryline Boval, J B Daniel

► **To cite this version:**

Daniel Sauvant, M P Letourneau-Montminy, Philippe Schmidely, Maryline Boval, J B Daniel. Use and misuse of meta-analysis in Animal Science. 2019. hal-02181827

HAL Id: hal-02181827

<https://hal.science/hal-02181827>

Preprint submitted on 12 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Use and misuse of meta-analysis in Animal Science

D.Sauvant¹, M.P.Letourneau-Montminy², P.Schmidely¹, M.Boval¹, C.Loncke¹, J.B.Daniel³

¹*Inra, AgroParisTech, Université Paris-Saclay, UMR Modélisation Systémique Appliquée aux Ruminants, 75005, Paris, France.*

²*Université Laval, Faculté des sciences de l'agriculture et de l'alimentation, Département des Sciences Animales 2425 rue de l'Agriculture, Québec, G1V 0A6*

³*Trouw Nutrition R&D, PO Box 299, 3800 AG, Amersfoort, the Netherlands*

Corresponding author: Daniel Sauvant, daniel.sauvant@agroparistech.fr

Short title: Meta-analysis in animal science

Abstract

In animal sciences the number of published meta-analyses increases with a rate of 15% per year highlighting an actual success. This current review focuses on the good practices and traps in the conduct of meta-analyses in animal sciences, nutrition in particular. The implementation of a meta-analysis is done in several phases after the definition of the study objectives. Clearly described rules and principles of traceability should be applied as soon as the publications are collected and selected with a target of meta-analysis. Then, the coding phase is essential because it determines the quality of the graphical and statistical interpretations of the database. Following this step, the study of the levels of orthogonality of factors and of the degree of data balance of the meta-design represents an essential phase to ensure the validity of statistical processing. The issue of the choice between fixed or random effect to study and to control heterogeneity is also discussed. It appears on the basis of several examples that this choice does not generally have any influence on the conclusions of a meta-analysis when the number of experiments is sufficient. Finally, reflections are presented on the potential interest of meta-analyses in the context of systemic approaches as well as to improve mechanistic modelling work.

Keywords: Modeling, Database, Meta-design, Random and Fixed effects, Nutrition

Introduction

To achieve better supported quantitative conclusions about a research question, researchers came up with the idea of grouping them together. Although a wonderful idea, the question of heterogeneity across experiments has in fact arisen. This concern appeared a long time ago in the medical field (Pearson K., 1904) and then in that of agricultural field experiments (Yates & Cochran, 1938). The term meta-analysis came more recently from medical studies (Glass, 1976). As showed in **Figure 1a** there was an exponential evolution of the number of publications that have applied meta-analyses in animal sciences with a progress of 15 % year⁻¹. This trend, which follows with a lag of 10-15 years the area of medical sciences (Sutton and Higgins, 2008) are likely to continue for several more years. This rapid evolution is

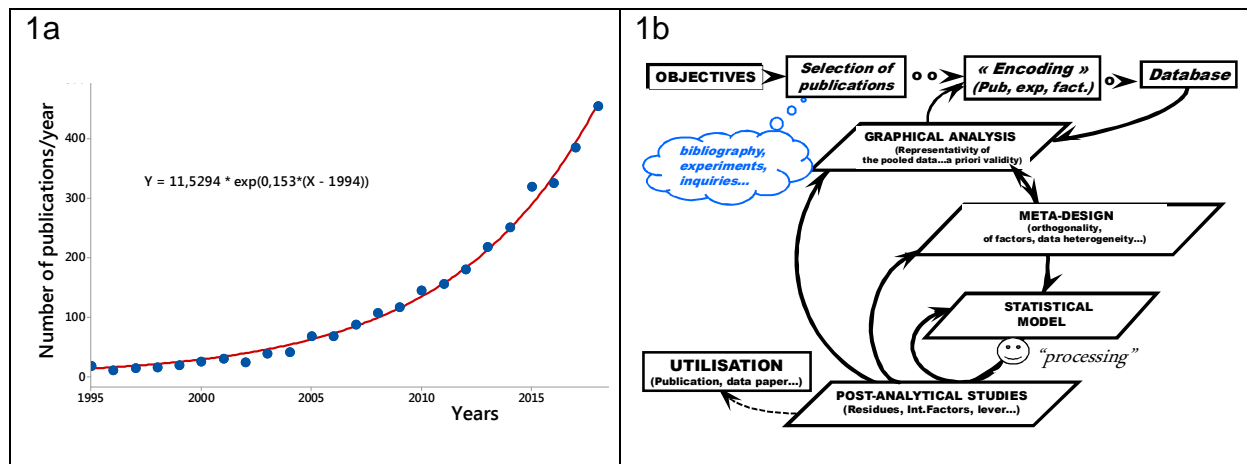
mainly due to the growing accumulation of experimental data per topic of interest (increasing numbers of publications / topic, and of measured data / publication...).

In animal science, meta-analysis has proven to be an efficient way to renew already published data by creating new empirical models allowing to progress in both understanding and prediction aspects. The progress is allowed by (1) the reduction of all biases and imprecision and (2) enlarging a priori the domain of validity of the model.

The publication of St-Pierre(2001) was an key step in the consideration, design and development of meta-analyses in physiology and animal nutrition. Indeed, it was the first publication, in a scientific journal, addressing the main "ins and outs" of meta-analysis and proposing a series of reflexions and conceptualizations relative to this issue. In particular it highlights the importance of splitting inter- and intra-experiment variations. Logically, this work has been cited numerous times in the field of Animal Science, so a stepping-stone for this tool. Subsequently, Sauvant et al. (2008) mainly focused the debate on the question of good practices to be applied in the meta-analysis in animal science. This second publication, carried out in collaboration with statisticians (N.St-Pierre and JJ.Daudin), insisted in particular on the good practices, on the graphical interpretation, on the choice between random and fixed effects of experiment, on the question of interfering factors and on that of the meta-design (see below).

The publications of St-Pierre (2001) and Sauvant et al. (2008) are now respectively 18 and 11 years old. Since these writings, many meta-analyses have been done and published (**Figure 1a**). Beyond these publications, meta-analyses have also proven their usefulness for updating feed unit systems. For example, the recent book "Feeding System for Ruminants" (INRA 2018) was constructed on hundreds of meta-analyses that allowed to propose more than 500 empirical equations calculated from more than 25 various databases. It seems important to try to take a profit on all these works to go a little further in the definition of good practices on the one hand and the limits of meta-analysis on the other hand. Our comments will be particularly based on the experience of the co-authors who published meta-analysis since early 2000 (see WOS...). Sauvant et al. (2008) proposed an heuristic cyclical generic approach with successive steps to conduct the meta-analyzes (**Figure 1b**). This heuristic approach is not questioned but it can be a bit updated thanks to different publications on the subject and serves as a basis for the plan. This approach is rather similar to that proposed in Agronomy by Philibert et al. (2012).

Figure 1: Evolution of the numbers of publications when crossing the key words 'Meta-analysis x Animal' in the Web of Science (1a). Graphical representation of the meta-analytic process, updated from Sauviant et al. (2008) (1b).



Specificities of meta-analysis in Animal Science

As indicated by St-Pierre (2001) and Sauviant et al. (2008), meta-analyses in animal sciences are generally quite different in terms of methods and objectives from meta-analyses conducted in the medical domain. However, few meta-analyses have also been applied in animal sciences following a similar approach that the one used in medicine (Phillips, 2005; Hillebrand, 2009; Srednicka-Tober et al., 2016).

Meta-analyses in Animal Science are mainly interested in relations between variables and are mainly aimed at predicting the average quantitative response (Y), within the experiment, to one or more independent quantitative explanatory/predicting variables ($X(s)$) such as $Y = f(X)$. In such variance-covariance analysis, one of the benefit is that from intra-experiment variance one seeks to extract a generic empirical model $Y = f(X)$, for example of a causal feeding practice quantified by X , which has been studied in a set of experiments. Under these conditions, the experiment effect corresponds to the variations between studies not taken into account by the covariables. In terms of design, we thus find a structure of repeated data (in intra-experiment) in space (experiments) which makes it possible to take advantage of statistical processing programs for repeated data in space. The fact that the covariable is systematically calculated intra-experiment constitutes also a difference with the medical field where the meta-regression concern basically the use of a covariable to explain the heterogeneity across studies (Borenstein et al., 2009).

First steps in meta-analyses

Selecting the publications.

One of the hallmarks of meta-analysis compared to conventional literature reviews is its comprehensiveness on a subject with an exhaustive collection of candidate publications based on a set of key words that are closely consistent with the objectives of the work. These key words can concern either certain factors or experiment or specific measurements. This harvest must start with the most generic keywords that are refined gradually to arrive at the list of publications to be eligible for the analysis. Through that phase the reduction of the number of candidate publications may be important. For instance, d'Alexis et al. (2014) conducted a meta-

analysis on mixed grazing on 9 publications after starting from an initial set of 8044 references, reduced to 117 eligible candidate with "mixed grazing" as a first filter. Candidate publication filtering is based on critical assessment of each of them, focusing on the detection of obvious important errors; data quality step (mainly based on the expertise of the analyst). Doubts will generally subsist for some publications and it may be wise to enter them and compare them to the rest of the data and decide from there to keep them or removing them. This visualization is indeed helpful as an ultimate quality screening. One important aspect during this selection step is to explicitly mention the reasons for the exclusion of the publications that are removed at the last steps of the selection process. This is good practice and allows the reader to understand why some candidate articles have not been included. In medical science proposals of flow diagrams to be applied, such as the Prisma one (<http://www.prisma-statement.org/>) have been made, they are globally consistent with our current description.

Data structure challenges.

The result of pooling publications is a table of data where rows represent treatments, while the columns consist of the measured variables and characteristics. One of the features of this dataset is to present numerous missing data. This seriously limit the possibility of using multi-variate statistics such as the principal component analysis (PCA). Therefore analyses must be performed by successive steps on small subset of independent variables, generally two by two to minimize the loss of information. Missing data are particularly numerous when the target is to focus on rarely measured characteristics or factors, due to difficulty or high cost of the measures. The presence of this missing information leads to a significant and variable information loss before being able to answer certain questions related to the objectives of the work. For instance in a recent meta-analysis focused on grazing behaviour of ruminants (Boval and Sauvant, 2019) an as exhaustive as possible database of 109 publications, 263 experiments(nexp) and 905 treatments (n) was gathered, the most measured behavioural item being the bite mass (documented in 64% of treatments).But, to study bite depth, the corresponding numbers were only 21, 74 and 225. Moreover, for testing the influences of variables such as sward height or herbage bulk density, the numbers of experiments and treatment available decrease even more(nexp=53 or 22, and n=126 or 69).

From a table of data to a database by encoding.

Once all the publication's data has been entered, we have a table of data that cannot be exploited as such. It is necessary to transform the initial dataset into an organized database suited for a meta-analysis process. Indeed, a preliminary step of coding the data is essential because, by this way, data are becoming accessible and useable. Such coding will also be involved in either graphical and/or statistical procedures (**Figure 1b**). Therefore this coding step is crucial to make a meta-analysis that will ultimately generate new reliable and generalizable knowledge. This coding can only be successful if not successful, only if the person doing the meta-analysis has true expertise on the subject and well aware of coding methods. To summarize, the scientific value and even the "art" of meta-analyses directly depends on the quality of this coding.

A first step, whatever the objective, is to code all the publications and all the experiments (or studies) carried out in each publication. As such this ~~experiment~~ code is complete but it contains ambiguity in the sense that it mixes experiments that

may have various objectives. Therefore, for an in-depth analysis, it is necessary to code each experimental objective in a separate column. Each of these codes corresponds to a factor of variation which can be studied in the meta-analysis. Such codes are without any ambiguity and, by this way, numerous columns will be gradually added. The list of these factors trace a quick and interesting portrait of what has been studied and vice versa from what has not been studied and could merit to be. This is an important step to assess the degree of validity of the empirical models proposed at the end of a meta-analytic process. These specific codes of factors are necessary to study the meta-design (see below). They can also be combined to model responses to various factors after having checked their mutual orthogonality (see below). Otherwise, according to the objective of the work it can be necessary to create new codes. For instance if the objective is to study the interaction between two factors A and B, it is necessary to create, beyond coding of A and B, a new code able to consider all data candidate to study the interaction $A \times B$. Directly based on data encoding, graphical examinations of the data are recommended all along the process of meta-analysis to enhance a global view as well as to identify specific relationships to be investigated. These aspects have already been developed by Sauviant et al. (2008).

The much talked about meta-design

Classically, experimentation conception involves a careful thinking of the experimental design to ensure the independence of the studied factors. In meta-analysis the structure of the data corresponds to a design which is a priori neither orthogonal (independent) nor balanced. This can lead to important statistical estimation problems. The major trap concerns collinearity between explicative variables, which may bias the interpretation of results, if not carefully considered. In particular one cannot estimate separately the effects that are more or less completely confounded with others. Therefore, a critical study of the meta-design is a key step in meta-analyses (**Figure 1b**). These aspects were already evoked by Sauviant et al. (2008), so only some new considerations or examples will be presented. To characterize the meta-design, several steps must take place before and after the statistical analyses.

Relations between qualitative factors and independent variables X

First, it must be checked that the experimental factors and covariates are independent before interpreting them. Misuses linked with this aspect were already evoked by Sauviant et al. (2008) with an example of a curvilinear intra-experiment relationship associated with a non-independence between X and the meta-design, other examples are considered further.

This situation can arise when there is continuous and categorical covariates to study. In that case, it is particularly important to make sure that variation of continuous X is similar between each category. As an example, a meta-analysis (Létourneau-Montminy et al., 2018) studied the effect of crude protein supply on daily water consumption in broilers of different age. However in this dataset, age was expressed as a categorical variable (0-21 days and 22-42 days) and because variation in crude protein were not similar between the two age categories, it was not possible to take into account the effect of age. Adding the effect of age, showed highly significant model, but not in concordance with bird response within the publications.

Relations between the quantitative independent variables X

When there is only one covariable, major aspects to be considered (histogram, study effect, lever effects...) were already listed by Sauvant et al. (2008). The complexity increases even more when two or more independent variables are considered together, especially if their interactions are also studied (see below). In such case, plotting independent variables against each other and quantifying their correlation is needed to assess degree of multi-collinearity, situation that may occur when two or more predictor variables in a regression model are redundant. This is quite usual in animal nutrition especially in monogastric receiving complete diet because of feed formulation practice that involves many ratios [e.g. amino acid, calcium (Ca) and phosphorus (P), dietary electrolyte balance].

In this case it is important to study carefully the orthogonality between the candidate factors and to estimate the intra-correlation between them. For example, Daniel et al. (2016) have pooled experiments on dairy cows dealing with either net energy or metabolisable protein supplies to model their productive responses and to test eventual interactions between both nutrients. In this work, the intra-experiment correlation between daily net energy supply and daily metabolisable protein supply was naturally high ($R^2=0.42$) due to simultaneous variation of feed intake within experiment. The consequence of this correlation is that the coefficient of the milk response attributed to protein and that attributed to energy cannot be interpreted independently. However, when the number of experiments on a specific subject is large, it is possible to reduce this correlation by selecting only the experiments with low correlation between the two factors of interest. As an example by selecting experiments with low variation in intake, the correlation between net energy and metabolisable protein supplies was reduced from $R^2 = 0.42$ to 0.13 (Daniel et al., 2017). In this case, the estimate of the effect of metabolisable protein supply on milk response approximate the "true effect".

Specificity of dynamic data

Publications with meta-analyses have so far mainly concerned static data. Nevertheless it is quite possible to apply them to dynamic data, that is, evolving over time. It was for instance the case of published kinetics such as lactation curves (Martin and Sauvant, 2002) or postprandial changes in rumen pH (Dragomir et al., 2010). A difficulty related to this type of approach is that the different kinetics taken into consideration are not consistent because of the different time measurements and intervals from one publication to another. It is therefore necessary to make a preliminary adjustment of the data in order to be able to conduct the analysis, the adjustment model can be continuous (Dragomir et al., 2010) or by segments when it makes sense with respect to the objective work (Martin and Sauvant, 2010).

Interactions between the effects of factors or covariables X

Studying the interactions across factors or covariables is an important challenge in meta-analysis. When a database contains only experiments designed according to simple factorial designs (2x2 or 2x3...), as stated earlier a specific column can be consecrated to code the interaction and the interpretation can be based either by analysis of variance or by two covariables representing measures on both factors. For instance such an approach has been applied to study the interactions between chemical and physical fibre in cattle (Sauvant and Yang, 2014), the interactions between P and Ca in broilers and pigs (Létourneau-Montminy et al., 2010, 2012) and

the interactions between leucine and valine in chicken (Zouaoui et al.,2019). Unfortunately this type of situation is fairly rare because within a topic the number of factorial useable experiments is generally insufficient. More generally the purpose is to study an interaction between factors from a database where no, or only few, experiments following a factorial design and several situations can be encountered. Often, but not systematically, the code of the publications can be used to help studying an interaction. For instance, Boval and Sauvant (2019) studied the marginal influences of sward height (SH) and herbage bulk density (HBD) on bite mass through two types of independent experiments focused on either SH ($n_{exp}= 51$, $n = 296$) or HBD ($n_{exp}= 15$, $n = 45$) impacts. To model the interaction, a set of 30 publications ($n = 339$), including not only these two types of experiments but also some other experiments with both data were selected. Otherwise, if the number of experiments is important and fairly balanced between two factors, an interaction can be extracted with a satisfying degree of precision. Thus Sauvant et al. (2011) pooled 158 experiments ($n=450$) focused on either influence of level of Dry Matter Intake (DMI) or on impact of proportion of concentrate(PCO) to predict CH₄ production including an interaction between DMI and PCO and quadratic terms for these 2 variables.

Responses of the dependent variable Y to independent X variables

These aspects were largely depicted by Sauvant et al. (2008). Briefly the analyst must examine graphically and then statistically the inter- and intra-experiment(or within-study) relationships between the predictor variables.

Interest of exploring a database from different angle: Y as a function of X and X as a function of Y

In some situations the question arises whether it could be possible for the same database to be valued for different purposes under various meta-designs, i.e. for studying both intra-experiment and inter-experiment variation. For example, calorimetric experiments have provided measurements for metabolisable energy (ME) and net energy ($NE = ME - \text{Heat Production} = k \times ME$, k being efficiency of use of ME to NE). From these measurements, numerous authors have proposed some empirical models that estimate NE from ME and some other models that estimate ME from NE. This was mostly performed by grouping all the available treatments without any distinction of heterogeneity factors (publication, experiment ...). As a result, there is some confusion in the interpretation of the results obtained. This is inconvenient since the animal maintenance requirement and the efficiency of ME to NE, derived from these equations and applied in the feeding unit systems, are sensitive to the type of approach. Thus it was suggested by Salah et al. (2015, 2016) and Sauvant et al. (2018) to use from a same database different approaches specific to the objective. If the aim is to estimate animal responses to ME supply (e.g. **Figure 2b**), then the equations should be adjusted for the effect of experiment in order to focus on the within-experiment variation, expected to be mainly driven by the "push effect" of the level of ME (assuming the data were properly coded, see section above). However, when the objective is to estimate animal requirement, within-experiment variation are, in most cases, of limited value. This is because within an experiment, animal factors, such as body weight, are made homogenous by the investigators, so that differences between treatments are only attributable to the experimental factors being under study (i.e effect of different level of ME). In contrast a large variability may be expected in animals factors between experiments, and this "pull type" variation is a

priori of great value to derive equations that can predict animal requirements (**Figure 2a**). For instance, from 87 calorimetry experiments (n=239 treatments) on lactating cows (n=187) and goats (n=52), three modelling approaches were compared:

-(1) an inter-experiment GLM procedure (**Figure 2a**):

$$ME/BW^{0.75} = SP + ME_m/BW^{0.75} + (1/k_i) (NE_{L\pm R})/BW^{0.75}$$

calculated across experiments (1 point = 1 experiment, ME_m is the ME maintenance requirement, NE_L is milk energy and R is the body retention/mobilisation of energy) within species SP as fixed factor without (Table 1, model 1a) or with (Table 1, model 1b) weighing each experiment as suggested by St-Pierre (2001). This model could be considered as a meta-regression in the sense of medical science.

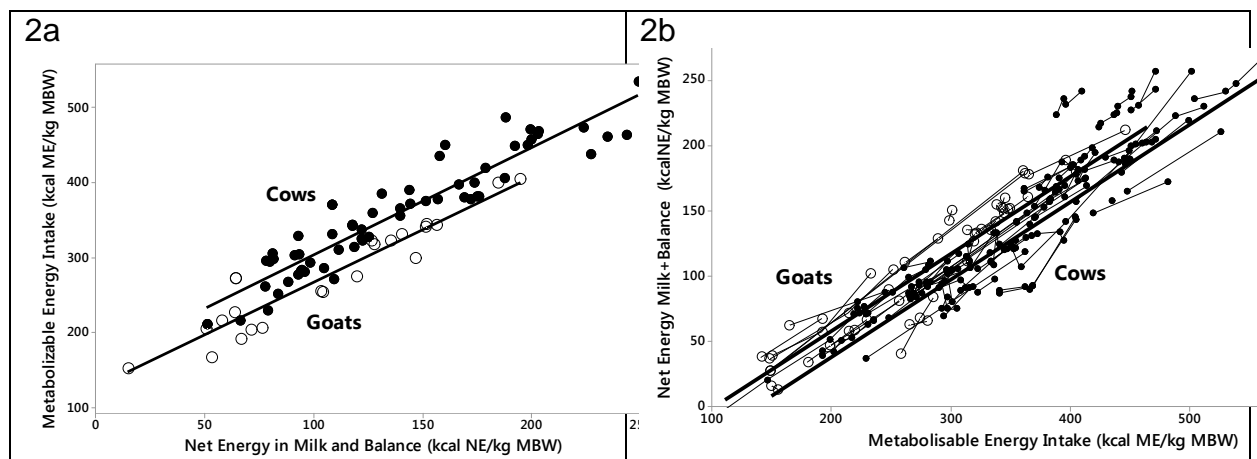
-(2) an intra-experiment GLM procedure (**Figure 2b**):

$$(NE_{L\pm R})/BW^{0.75} = a + k_i(ME)/BW^{0.75}$$

adjusted for the fixed effect of experiment (1 point = 1 treatment) nested by species.

-(3) a Mixed procedure (covariance matrix VC) similar to model (2a) except for the effect of experiment considered in this case random instead of fixed.

Figure 2: (2a) Inter experiment relationship between ME intake and the sum of net energy partitioned into milk and to/from the body (with 1 point = average of all of the treatment of one experiment) (2b) intra-experiment relationship between the sum of net energy partitioned into milk and to/from the body with ME intake (with 1 point = 1 treatment). Both relationships were obtained using the same database consisting of lactating cows and goats.



The **Table 1** summarizes the estimated maintenance ($ME_m/BW^{0.75}$) and efficiency from ME to NE (k_i) obtained from the 4 different models tested with data from lactating cows and goats. From this data, it is evident that the choice of the model influences these estimations. In particular the estimation of maintenance requirements and ME efficiency are higher with models (1a and 1b) compared to models (2 and 3). Moreover the best precision, assessed with RMSE, is achieved for the models which adjust for the effect of experiments (2 and 3). Otherwise, if data from dry cows and dry goats are added to this lactating dataset (as a mean to include feeding practice closer to maintenance), then the same models give different estimations (unpublished data). The diversity of these results highlight the importance of the approach taken for estimating these key parameters of energy unit systems.

This also demonstrates that the interpretation of such results need to precisely consider the type of approach used.

Table 1. Comparative estimations of maintenance requirements and efficiency of ME into NE of milk and body reserves (kl).

Model	Maintenance for cows & goats (\pm SE) (kcal ME/ BW ^{0.75})	Efficiency of ME to NE kl = NE/ME (\pm SE)	RMSE
(1a) Inter-experiment (n _{exp} =85)	162.5 & 125.7 (\pm 7.61)	0.705 (= 1/1.419 \pm 0.057)	26.0 (kcal ME/ BW ^{0.75})
(1b) Inter-experiment (weighted by n ^{1/2})	166.6 & 130.6 (\pm 6.80)	0.782 (= 1/1.371 \pm 0.048)	21.5 (kcal ME/ BW ^{0.75})
(2) Intra-experiment (fixed)	149.9 & 116.3 (\pm 5.0)	0.660 (\pm 0.016)	6.9 (kcal NE/ BW ^{0.75})
(3) Intra-experiment (random)	147.0 & 114.2 (\pm 7.6)	0.654 (\pm 0.014)	6.9 (kcal NE/ BW ^{0.75})

Evaluation of models obtained from meta-analysis

First it must be recall that the rigorousness of the meta-analytic approach is essential insofar as certain aspects may be the result of subjective choices or questionable "expert statements". In this context, a first essential point is to provide clearly the justification of the successive choices made during the analysis. Another aspect concerns the traceability of the whole approach, on this point of view some software (i.e. Minitab, SAS, R...) systematically report the work history. Beyond this question on the rigor of the approach is that of the repeatability of the work according to the person or the team that performs it. In fact, depending on the degree of knowledge of the subject and the expertise of the analyst, most of the major results will be similar, but in some minor aspects the approaches may diverge depending on the studies.

A priori the validation of a model resulting from a meta-analysis has to be conducted on an exogenous dataset that did not contribute to the model calibration. However this is a tricky approach since such comparison must be done on a dataset with very similar characteristics in term of experimental factors (i.e. experiment varying energy level) and range of data studied (i.e. range of energy variation). Obviously this approach is not possible for data rarely measured for reasons of cost or technical difficulties. In fact, if the principle of exhaustive data collection is applied, there is no other exogenous data available. In such a context it seems that the assessment of validity of the model resides first in the precise study of the representativeness of the collected factors and data used for the meta-analysis are:

- the factors present in the database representative of what is observed in practice in terms of existence, intensity, range of actions....? Is it possible to study the interaction of certain factors which are known to interact in field conditions ..?. All these aspects are frequently neglected despite they are strong determinants of validity of a meta-analytic work.
- for each factor, the data representative of the corresponding ones which are known in practice (range, frequency, statistical distribution...) ? How are they situated in function of the possible and plausible areas for each variable ?

The assumptions of nature of experimental heterogeneity

The debate between the choice of considering heterogeneity across experiments with either a random or fixed effects is not new and many papers have addressed this issue, particularly in the medical field where this choice can lead to opposite decisions, such as authorizing or not a drug to be marketed (Thompson, 1994, Petitti, 2001, Higgins et al., 2003, Borenstein et al. , 2007). This debate reflects the fact that, even if global principles are admitted, it appeared there is no objective (and unambiguous) method universally accepted for choosing between fixed and random effects models. Moreover, it must be indicated that in medical science Higgins et al. (2003) proposed to calculate an index of inter-experiments heterogeneity (I^2 index) which is an index of dispersion between studies for the measured item. This method, however, do not seem to be consensual and no similar test have been proposed sofar in Animal Science.

Random effect

In this case the heterogeneity between experiments is considered as the result of a random sampling within a large population. Experiments are therefore assumed to be independent and therefore they cannot share a priori a common objective (see unambiguous coding and part 2.1). The objective is essentially to control and model the heterogeneous variance and to achieve a prediction taking into account the intra- but also at least partly the inter-experiment variability. This last random variation induced by studies can largely increase the confidence interval (CI) values of parameters and therefore reduce the value of the regression for prediction purposes. The variations induced by this heterogeneity are supposed to follow a known random distribution, in preference a Gaussian one (e.g. normal). The ambition is to be able to apply the results obtained to predict or infer any future experiment from the entire population (whole area of inference). In such a context it is logical to take into account not only the random residuals of the regression but also, at least partly, the variations across studies assumed as random.

This principle has inspired many authors of publishing meta-analytical work, in which the experiments were considered random. However, in most cases, this choice has not been justified, and the only argument for this choice was to refer to Saintt-Pierre (2001) publication, which is by construction a random one.

On the other hand, it seems important to remember that the choice of a random factor leads to renounce, at least momentarily, to any explanatory examination of the experimental heterogeneity. Additionally, one can also wonder about the meaning to give to an interaction between an intra-covariate which is a fixed factor with a randomized experiment effect ? Finally, it is good to keep in mind that somehow the choice of a random effect is easy for the analyst given that there is no more question about experimental heterogeneity.

Fixed effect

Conceptually the choice of a fixed effect is when it is wanted to have both a mean prediction and to understand and interpret the causes of the heterogeneity across

experiments. It is close to the statement of Van Houwelingen et al. (2002) and from other comparable ones (Greenland, 1987; Berlin, 1995): « In the case of substantial heterogeneity between the studies, it is the researcher's duty to explore possible causes of heterogeneity ». Practically several contexts can justify the choice of a fixed effect:

Case of a "homogeneous" experimental context. Conceptually this case is close to the situation of the "single common effect model" described in medicine (Borenstein et al., 2010). In this situation the collected experiments share a common objective and factor and they are carried out in comparable working contexts. It is then assumed that there is no important heterogeneity on this aspect. A questionable aspect is which situation can be considered as a "no important heterogeneity" ? The goal can be, for example, to have an estimate of the average impact of the level of dietary protein supply on a given type of animal performance by selecting experiments that were all designed to study this aspect. In this case, the variation studied and tested is assumed to come a priori essentially from a sampling variability between the tests and the choice of a fixed effect is then recommended. In this context, the fixed effect is considered to be more powerful and precise since it relies essentially on the intra-experiment variability. For example, in the recent Systali project (INRA 2018) specific codings of experiments having the same objective within the same database "Bovidig" allowed to split and model the digestive influences in cattle through specific dietary factors such as protein, starch, cell wall, fatty acids...(Sauvant and Nozière, 2016). However, under this assumption, the effect can be applied only to comparable data or in the same overall context (inference). Thus obviously, the quality of inference would be better if the levels in the study represent all possible levels of the factor encountered in practice, or at least all levels about which inference is to be made (Littell et al., 1996).

Case of a large and partly explicable experimental heterogeneity:

- *Creation of subgroups.* If within a set of experiments major causes of inter-experiment heterogeneity are identified and sufficiently represented, it is suggested to return to the situation above by making subgroups of experiments having common contexts and low internal variability to apply fixed-effect to the subgroups. In this case the variance inter-experiment-intra-group is used as residual to test the fixed effect, through a nested model, and then it is considered as random for the F test. For example in a meta-analysis studying interest of mixed grazing, D'Alexis et al. (2014) proposed, after studying the database, the following nested structure: climate (1 DF) < publications intra-climate (7 DF) < replications within publication (8 DF) < types of associated animals (10 DF) < treatments. In such a situation the meta-design has to be carefully considered, particularly if the objective is to study the interactions between the subgroups and the covariables.

- *Creation of an inter-experiment covariable.* In a similar context, if the inter-experiments variance is largely explained by a covariable, it corresponds a priori to a situation comparable to that of meta-regressions in medicine (Borenstein et al., 2009). For example, we have carried out a meta-analysis of 169 experiments (420 treatments) focused on dairy cows milk yield ($MY=30.5 \pm 7.0$ kg/d) responses to dry matter concentrate supply ($DMlc =9.7\pm 0.16$ kg/d). Variations across experiments are important (Figure 4a) and largely explained by the potential milk yield (MY_{pot}) of each experiment (= common value between the treatments within a given experiment

= 30.8±0.3 kg/d) (Sauvant and Daniel unp.)). In this example the inter-experimental effect on MY is tremendous ($R^2=0.95$) and MYpot is also highly influencing ($R^2=0.93$). In such a situation, the analyst has a priori the choice between considering an experience effect or replacing it with a covariable equal to the production potential of each experiment. However in this case to interpret the coefficients of regression, it is important to ensure, as stated before, that the covariable is not linked with the independent variables (DMIco and DMIco²). Several statistical models were fitted on this dataset (Table 2). When the effects of experiments (fixed (1a) or random (1b)) were replaced by the MY potential of each experiment (model (2)), the intercept became non-significant, which is logical, and the coefficients of DMIco and DMIco² were largely different from the corresponding values of fixed (1a) and random (1b) effect assumptions, these last values having a concrete sense in term of animal response and diet formulation (Faverdin et al., 2018). Moreover the coefficient of DMIco² became non-significant and, based on the RMSE, the precision of the model, was less. These large modifications in the values of these two coefficients are mainly due to the fact that there is a significant positive inter-experiment relationship between the potential milk yield and level of concentrate supply ($R^2=0.40$). This relation is logical in the sense that cows having a higher potential must receive more concentrate to meet their energy requirements. Finally, when model(2) is compared to model (1) the adjusted values are globally similar ($R^2=0.98$ and not different from $Y=X$) with a significant remaining influence of the experiments not tackled by the covariable MYpot, the residuals are less linked between (1a) and (2) ($R^2=0.47$) with also a significant influence of the experiments. Therefore it is important to carefully check the data to avoid the consequences of a correlation between a covariable candidate for representing the experimental heterogeneity and the intra-experiment covariables.

Table 2: Results of intra-experiment fitting milk yield response of dairy cows to concentrate supply (\pm SE)

Model	Cste	Coef. of MYpot	Coef. of DMIco	Coef. of DMIco ²	RMSE
Fixed (1a)	24.18 (± 0.420)		0.827 (± 0.086)	-0.016 (± 0.004)	1.10
Random (1b)	23.90 (± 0.618)		0.870 (± 0.086)	-0.017 (± 0.004)	1.10
MYpot cov (2)		0.899 (± 0.011)	0.296 (± 0.070)	-0.001 ns (± 0.003)	1.60
Fixed (3)	30.6 (± 0.085)		0.499 (± 0.022)	-0.023 (± 0.010)	1.12
Fixed (4)	30.02 (± 0.077)		0.526 (± 0.022)	-0.018 (± 0.004)	1.06

Another important aspect concerns the mode of expression of the independent variables X. Indeed these variables can be expressed as such (models (1) and (2)) or with respect to a value of practical interest. For example, we adjusted the data with a fixed effect but according to two other modes of expression of DMIco: model (3) by centering it on the mean value within each experiment and model (4) by calculating the difference between DMIco and its adjusted value corresponding to the energy balance = 0, a meaningful nutritional pivot (Faverdin et al., 2018). Table 2 gives the

value of the two adjustments. Intercepts and coefficients of regressions are different from those of the model (1a) and these differences can be explained. For instance the marginal response of MY to concentrate is high (0.83 g/g DMIco) when DMIco is close to 0, but much less (0.53 g/g DMIco) when energy balance = 0. Moreover, between models (3) and (4) the adjusted (or predicted) values are similar ($R^2 > 0.999$, close to $Y=X$) and the residuals are highly correlated ($R^2 = 0.91$). Both criteria being closely related ($R^2 = 0.91$) with the corresponding values of model (1a), ($R^2 > 0.998$ and $R^2 > 0.84$) and model (1b) ($R^2 > 0.998$ and $R^2 > 0.89$). These results show the importance of having a close consistency between the future practical use of a model and the mode of expression of independent variables. In this way models issued from published meta-analysis cannot be used without a careful consideration of their process of construction.

Case of a large but non-easily explicable experimental heterogeneity.

Several reasons can lead to this type of situation: (1) if experimental variations are caused by well-known causes, the meta-design can be too imbalanced and/or non-orthogonal to take these effects in the analysis without risk (2) if the meta-analysis is performed on a dataset of experiments having diverse objectives with no possibility of doing sub groups of experiments with a common target (3) if measurement methods across studies are not sufficiently standardized etc....

Misuse linked with an heterogeneous situation that is not easily explicable and treated as a fixed effect: it is generally considered that, when the inter-test heterogeneity of the effects is important, and not easily explicable, and treated as a fixed effect, the confidence intervals of the intercept are too narrow (illustrative examples are given below in the text). Another point of view (Saint Pierre, 2001) is that an effect can be considered fixed if the levels in the study represent all possible levels of the factor, or at least all levels about which inference is to be made (Littell et al., 1996).

Comparisons between random and fixed effects

In fact, according to the objective, the choice of a random or fixed effect can be justified. For example, in the case of a Latin square associating n diets, n periods and n animals, it is possible to decide that the animal effect is random if the goal is only to control the variance to make a priori the test on the diets more powerful. In contrast, if it is wanted to understand differences between animals, for instance for phenotyping, a fixed effect could be considered.

Otherwise it must be recalled that, the inference range for the fixed model is strictly limited to those studies that are part of the regression while the inference of a random model is a priori applicable to a larger population. To give ideas of the differences to be expected in practice from the consequences of the choice between a fixed or random effect, it seems simpler to briefly describe some examples.

-Example 1: meta-analysis with numerous heterogeneous experiments.

In the published comparisons between the two types of effects with numerous experiments it has been pointed out several times that the intra-experiment regressions and the adjusted values of the parameters were very similar (see annexed documents of Loncke et al., 2015, and of Daniel et al., 2016). Moreover, the results presented in Tables (models (2) vs model (3)) and Table 2 (models (1a) vs (1b)) confirm that when the number of experiments is important (about > 30) the

intra-experiment regressions, the adjusted values and the residuals are very close between fixed and random effects assumptions. From all these examples treating numerous experiments, CIs were systematically more important in the case of the random effect for the intercept, in contrast for the coefficients of intra-experiment regression values of CIs were almost the same with random assumption being equal or slightly higher.

- Example 2: further study of the experiment effect in the dataset of Saint Pierre (2001)

Saint Pierre (2001) did a comparison between fixed and random effects on a theoretical data set of 20 studies (108 treatments) where effects of study on intercept and slope were correlated. To go a bit further in this interesting work we have reconsidered his data set by comparing 4 approaches:

- (1) Separate individual fittings of the 20 experiments by 20 linear regressions (IND).
- (2) GLM procedure (with experiment effect as fixed effect) including interaction between experiments and covariable X (GLM) with a fixed effect.
- (3) Mixed model with study effect as random effect with the hypothesis of "Variance component" (VC) structure of covariance matrix including interaction between studies and covariate as random effect.
- (4) id. than (3) with the "unstructured" structure of covariance matrix (UN).

Table 3: Comparison between the major parameters with the fixed and mixed models:

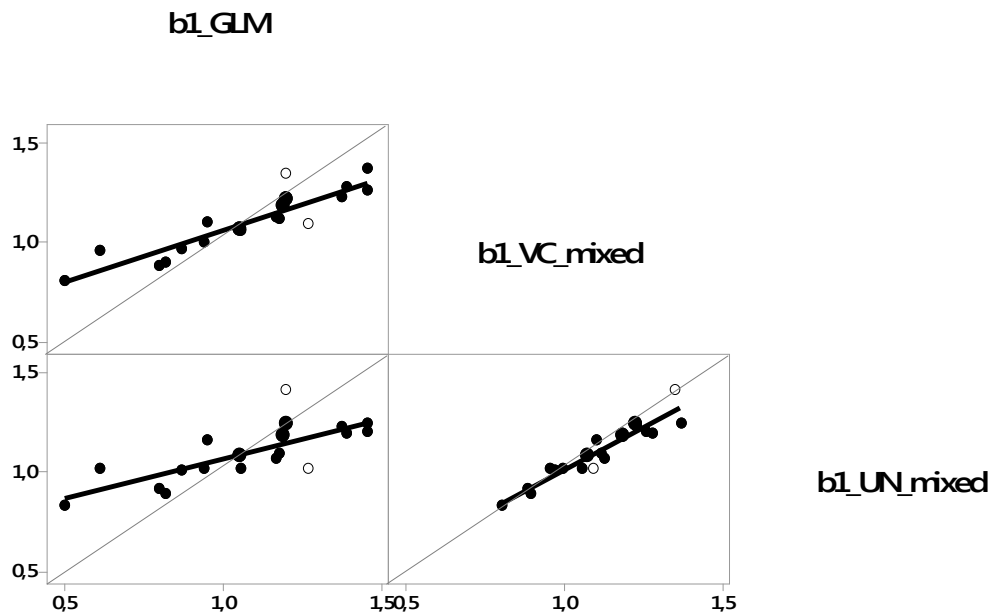
	Model 1 (IND)	Model 2 (GLM)	Model 3 (Mixed type=VC)	Model 4 (Mixed type=UN)
Intercept	-0.469	-0.469 (0.264)	-0.588 (0.606)	-0.608 (0.563)
Slope	1.0797	1.0797 (0.047)	1.0995 (0.061)	1.0974 (0.056)
RMSE	NA	0.50	0.49	0.50

¹Means (standard error)

The major aspects concerning this comparison are:

- Slopes and intercepts values of experiments are exactly the same for (1) and (2).
- The mean intra-experiment slopes are close for the models (2), (3) and (4), particularly for the two mixed ones (**Table 3**).
- The confidence intervals (CI) are larger for the mixed models: more than x2 for the intercepts but only x1.2 to 1.3 for the slopes (**Table 3**). This confirms the above mentioned observations in examples 1.
- Compared to IND and GLM, the slopes per experiment are different in the sense that the range of them is narrower for both random models (**Figure 3**). So the random effect assumption tend to reduce the range of intra-experiment slopes compared to the true intra-experiment values (canalisation effect). A major differences between fixed and mixed procedures concerns the experiments with 2 treatments with assumed values very different from the true one (IND) for mixed.
- Results obtained with both random hypotheses are very similar, may be because the dataset was not conceived to show differences between VC and UN.

Figure 3: Comparisons of values of the intra-experiment slopes (b1) with the dataset of SaintPierre (2001)



(1) open circles concern experiments with 2 treatments, large black circles experiment having 7, 8 or 9 treatments.

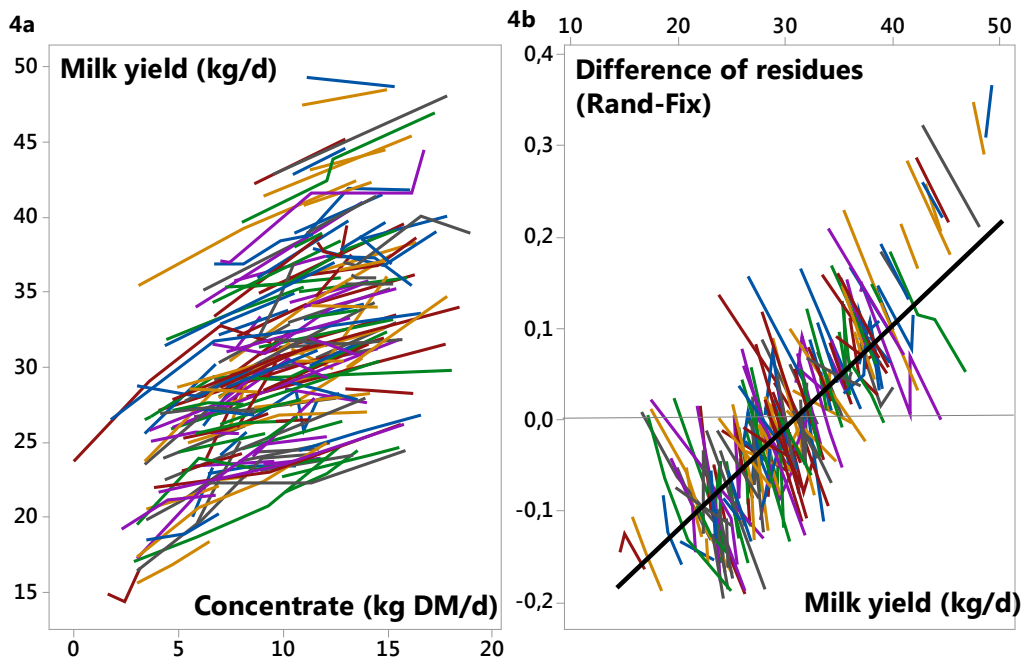
- Intercepts of experiments are mutually well correlated across the models ($R^2 > 0.87$) with a relationship not different from $Y=X$.
- No relationship appeared for fixed effect between intercept and slope while the random ones show a positive correlations ($R^2 = 0.24$ for VC and $R^2 = 0.30$ for UN). This is quite consistent with the way to construct the database assuming $R^2 = 0.25$ between slope and intercept values (St Pierre, 2001).
- Adjusted, or predicted, values of treatments are exactly the same whatever the model with $R^2 > 0.999$.
- Residuals of treatments are positively and well correlated ($R^2 = 0.84$ and 0.90 between GLM and VC and UN, and $R^2 = 0.98$ between VC and UN) and the relationships did not differ from $Y=X$. The differences between the residuals of random vs fixed effects hypotheses are largely explained by the lever effects (HI) attached to each treatment .
- Averaged values of residuals of each experiment are equal to 0 in the case of GLM while they are equal to 0.0019 ± 0.031 and 0.0013 ± 0.030 for the mixed model (VC and UN respectively). Moreover, whatever the mixed model considered, the averaged value or residual for each experiment is largely explained by the averaged value of Y and the number of treatments per experiment, the slope being much higher for 2 treatments/experiment (**Figure 5a**). Therefore, a part of the differences between studies of Y is recovered in the residual variance for mixed procedures.
- Globally the HI of treatments are positively correlated ($R^2 = 0.62$ and 0.46 between GLM and VC and UN, and 0.96 between VC and UN)

-Example 3: experiment effects and post analytical study in a practical situation

As the example of St-Pierre was a theoretical one we did a similar approach with the above example of dairy cows milk yield (MY) responses to concentrate supply (**Figure 4a**, $n_{exp} = 160$, $n = 420$) with the comparison between models 1a (fixed) and 1b (random, VC) of **Table 2**. In this case the mean slopes of individual experiments were highly correlated ($R^2 = 0.99$) and the relation was not different from $Y=X$.

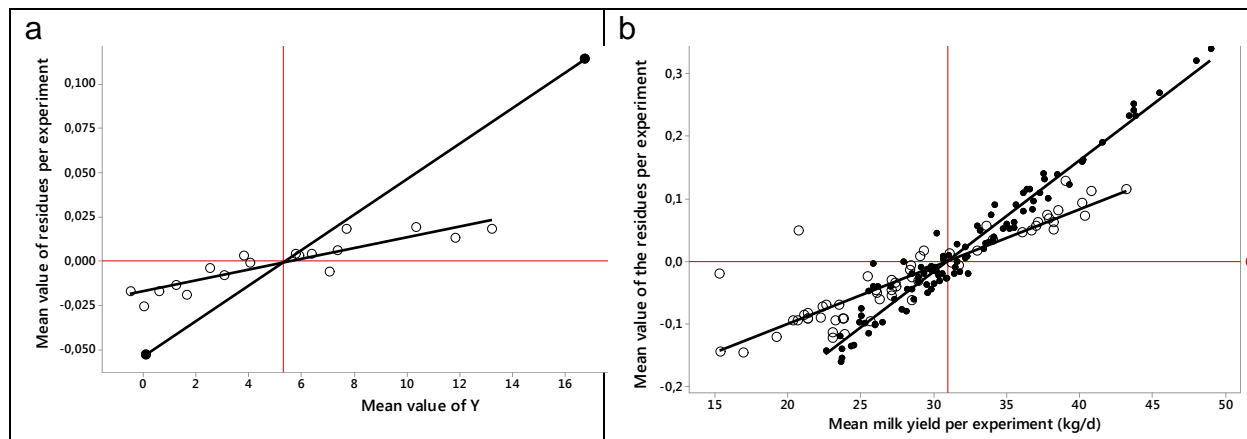
Residuals were highly correlated ($R^2=0.986$) and the global relation was not different from $Y=X$. The differences between residuals of random and fixed effects are related with MY, positively inter-experiment and negatively intra-experiment (**Figure 4b**).

Figure 4: Responses of milk yield (MY) to concentrate supply (4a) and relationships between MY and difference between residuals of random - fixed models (4b).



When residuals are averaged for each experiment, for both models the distribution of the mean residual standard errors for the different studies were not symmetric (χ^2 distribution), values were similar (0.91 ± 0.62 kg milk) and highly correlated ($R^2=0.996$ around $Y=X$). Consequently, if the analyst wants to remove experiments because of a high averaged residual value (Sauvant et al., 2008), these experiments would be the same whatever the type of effect. The analyst could also use these averaged residuals to weight the experiments in a re-run process to assess the stability of the results. Otherwise, with GLM (model 1a) residuals within all the experiments were centered on 0 while for the mixed procedure (model 1b), as in example 2, it remains an effect of experience because averaged values were different from 0 (0.004 ± 0.096). Practically, these variations across experiments were not linked with the averaged residuals mentioned just above, but as expected from **Figure 4b**, they are largely explained by the mean MY and by the number of treatments per experiment (**Figure 5b**). These results are quite consistent with what was observed just above on the example of Saint Pierre (**Figure 5a**). In both **Figures 5a** and **5b** the mean dependent variable has a stronger influence on the mean residual values for experiment that contains only 2 treatments.

Figure 5: Influence of the mean value of the dependent variable on the averaged values of the residuals per experiment in a random model (VC) for the examples 2 (5a) and 3 (5b). Black circles are experiments with only 2 treatments, white circles experiments with more than 2 treatments.



The knowledge of lever effects (HI) per treatment can be used to point at influent treatments (Sauvant et al., 2008) and are a good complement for the analysis of treatment with high residual (and thus potential candidate for outlier).. As seen above in the example of Saint Pierre (2001), HI values are much higher for the GLM than the mixed model, moreover they are much more influenced by the number of treatments per experiment with fixed compared with mixed model. When HI values from the fixed model are multiplied by the number of treatments, the relationship between these adjusted HI value and the one obtained from the random model are significantly related ($R^2=0.715$). However, there is a close intra-experiment regression between the two HI values: $HI_{fixed} = 1.00 + 1.98 HI_{mixed}$ ($R^2 = 0.873$, $RMSE = 0.012$) showing that, independently from the effect of the ratio treatment/experiment, treatments with highest HI values are very similar in both models. In addition to the number of treatment (in fixed only), the independent variable DMI_{co} had a major influence on HI in the two models. Highest HI values were found for treatments with either a low DMI_{co} (<5kg/d) or a high DMI_{co} (>15kg/d).

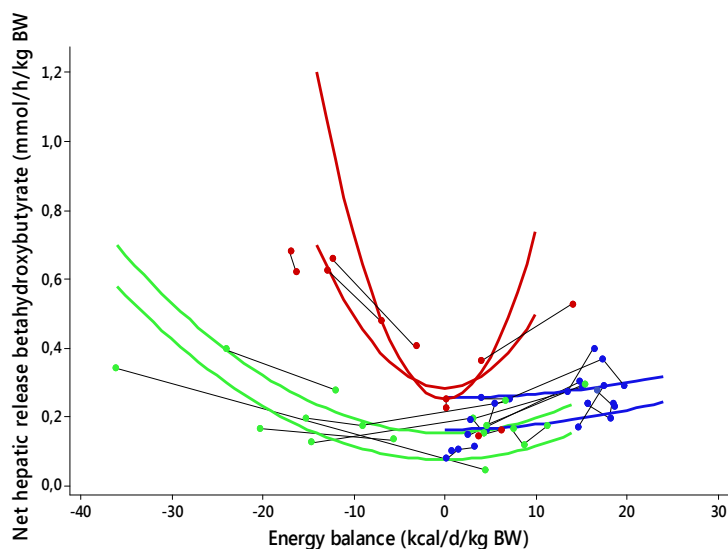
When HI values are averaged per experiment to detect those having important influence on the regression, it appeared that for both models the number of treatments has an influence but it is much more marked and negatively ranked according to the number of treatments/experiment for the fixed effect. When averaged HI values per experiment were adjusted by the number of treatments, the major factor was the standard deviation of the level of DMI_{co} per experiment (3.0 ± 1.6 kg DMI_{co}/d) showing that the positive influence of the range of DMI_{co} per experiment on the coefficient of MY response to concentrate. When HI are corrected from DMI_{co} impacts, there is a very close relationship between both corrected values: $corHI_{mix} = -0.033 + 0.97 CorHI_{glm}$ ($R^2=0.999$). That means that ultimately both effects led exactly to the same conclusion to eventually remove aberrant treatments or experiments on the basis of low/high residuals and HI.

-Example 4: case with a curvilinear response and an unbalanced design.

Loncke et al. (2015) have studied of the flux of β -hydroxybutyrate from the liver in ruminants as a function of energy balance. The database pooled 6 experiments with dairy cows ($n=12$), 9 with growing ruminants ($n=22$) and 8 with non-productive ruminants fed at maintenance ($n=20$). The intra experiment response was curvilinear

and it appeared that there are some differences between the two types of models (**Figure 6**). The major one concerns the dairy cows with predicted values much higher for random model when energy balance was either low or high. Indeed, the differences between GLM and Mixed models were mostly apparent when the interaction between the effect of energy balance and the animal physiological status (lactation, growing, maintenance) was taken into account. When this interaction was removed, the models were statistically non different. One explanation of the differences between GLM and mixed models in the present case could be the number of observations in each physiological status which is unbalanced (**Figure 6**). Moreover, the number of available data in particular in lactation (when the cow mobilized energy) is low.

Figure 6: Observed and adjusted net hepatic release of β -hydroxybutyrate relative to energy balance for dairy cows (red) growing cattle (blue) and maintenance (green). The dotted line represent the adjustment with the fixed model and continuous line represent the adjustment with the random model.



Interfering factors

The issue of interfering factors (IF) is also sometimes an object of debate (Sauvant et al., 2008). Originally, the research for interfering factors consisted in investigating whether qualitative factors, or continuous but incomplete variables (because of missing values in the dataset), that are, for this reason, not incorporated into the adjusted statistical model, could explain a significant part of the variation in the data. More precisely, when the better independent X variable(s), which can be considered as the primer driver(s) has been found, based on statistical (and eventually on mechanistic) criteria, it is advisable to check if any secondary variation factors or variables could influence the response equation. However this approach can also be made for complete but a priori not essential parameters. The study of the IF can be carried out at various levels (1) on the LSMEANS (2) on slopes of the intra-experiment relationships and (3) in the residual variation.

A systematic study of interfering factors allows to have a more complete knowledge of the treated data, some examples were published in the literature (Loncke et al.

2009 and 2015; Agastin et al. 2014). In a meta-analysis (Loncke et al.,2015) using published data compiled from the FLORA database (FLuxes of nutrients across Organs and tissues in Ruminant Animals; Vernet and Ortigues-Marty, 2006),response equations were proposed that predict hepatic uptake or release of ketogenic nutrients to their supply to the liver. The study of interfering factors showed that the general response equations can be significantly affected by another nutrients or by the diet characteristics. For example, the net hepatic uptake of acetate could be predicted by its net portal appearance, but it was shown that the within-study slopes significantly decreased when net portal appearance of propionate increased. On the same database, they showed that the hepatic release of glucose was related to the net portal appearance of nitrogen (Loncke et al., 2019), but the average level of glucose release (LSMeans) was also positively and logically influenced by the dietary starch concentration.

In general, it is necessary to consider all factors and variables as potential interfering factors, and to test their influence on the parameters (LSMeans, residuals and within-study slopes). To give an example, Loncke et al. (2009, 2015, 2019) tested on average 50 potential interfering factors for each equations published. This type of analysis gives more information about the equations proposed, by identifying if the main independent variables have consistent effects across various scenarios or if those effects are function of specific conditions. This strengthens the importance of completeness when building a dataset. Indeed some variables that may seem secondary at first may be relevant for such test. It is therefore highly advisable to enter all reported variables from a selected publication into the database.

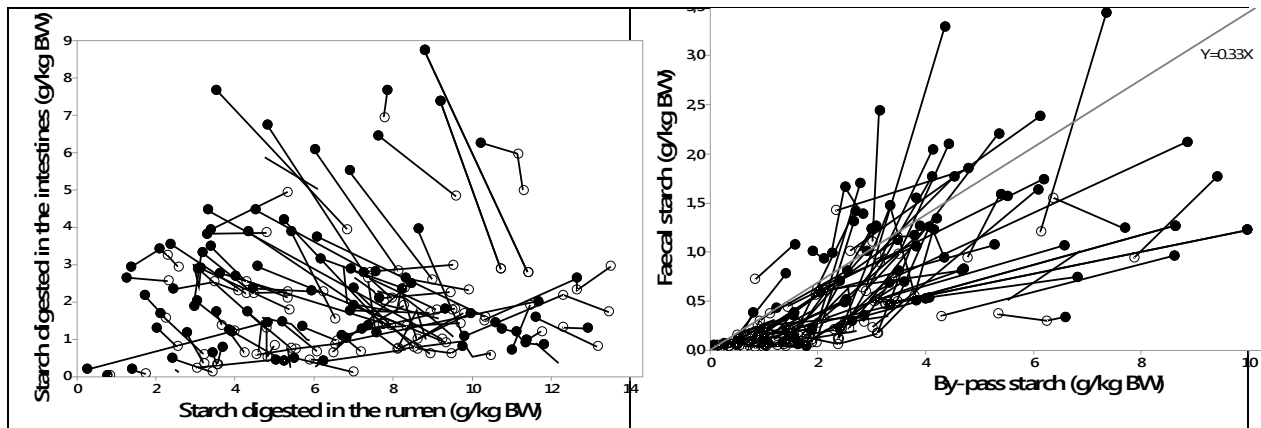
Meta-analysis and systemic approach

The modeling of biological systems is a difficult task as these systems are complex. Thus a challenge is to represent the main properties emerging from such system into a model. For that it is first necessary to know as precisely as possible the structure of such systems and the relationships between its elements and also its various levels of organization. In this context meta-analyses can be applied to understand the places and roles of the multiple scales within the spatio-temporal organization of systems. They can thus make it possible to highlight the key relationships associating the various elements and scales as well as their diversity of situations and their main factors of variation. Thus it can be particularly useful to help understanding emerging properties of complex systems. Such examples remained fairly rare despite their interest.

Offner and Sauvant (2004) have focused on the digestive system of starch in ruminants by pooling into a database a set of experiments exclusively focused on factors of variation of digestive flows of starch in ruminants. In this digestive system characterized by 3 successive compartments (rumen, small intestine, and hindgut),the meta-analyses performed have allowed to quantify the mean flows of starch between these 3 organs and thus the digestive partition of starch within each of them. Thus, **Figure 7a** presents the substitution between starch digested in the rumen (in g/kg Body Weight), providing VFA, and starch digested in the intestines, providing mainly glucose. This Figure highlights the importance of controlling ruminal starch degradability to control its digestive partition (Sauvant and Nozière, 2016). In addition **Figure 7b** presents the influence of flow of by-pass starch on starch losses

through feces which can be non negligible. These two figures also emphasize the large heterogeneity of situations across the diversity of the experiments. All the intra-experiment regressions on this starch digestive system were included in the recent INRA 2018 system.

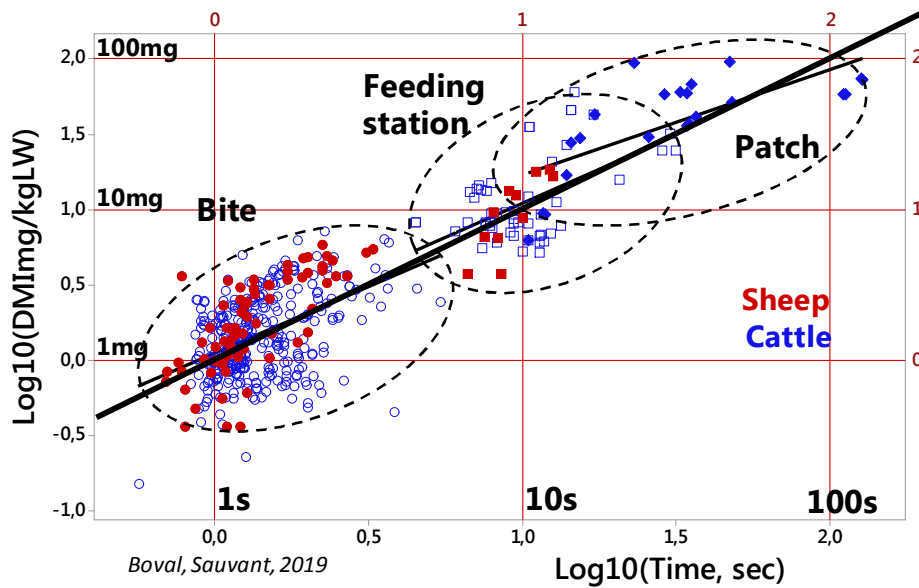
Figure 7: Substitution between Starch digested in the rumen and in the intestines (7a) and influence of by-pass starch on starch losses in feces (7b).



(1) open circles are treatments presented as quickly degrading starch and black circles are with slowly degrading starch.

Otherwise, meta-analysis can help to understand relationships between several levels of organization. Thus, in a recent meta-analysis focused on grazing behaviour, we studied the relationships between 3 spatio-temporal scales to better understand the organization of the short-term feeding and physical behavior of grazing ruminants; the 'bite', the 'feeding station' (FS = set of bites without moving front legs) and the 'feeding patch' (set of successive FS). A database including experiments with data at the 3 scales was created and interpreted accordingly (**Figure 8**). The meta-analysis shows that spatio-temporal areas of these 3 behavioral events are partly overlapping and follow globally the same simple allometric relationship ($\log_{10} \text{DMI} = \log_{10} \text{Time duration}$), moreover within each level the response to a given factor such as the sward height tends also to follow this same allometric law (unpublished data). Moreover it appeared that sheep (close figures) and cattle (open figures) can be differentiated at the bite level.

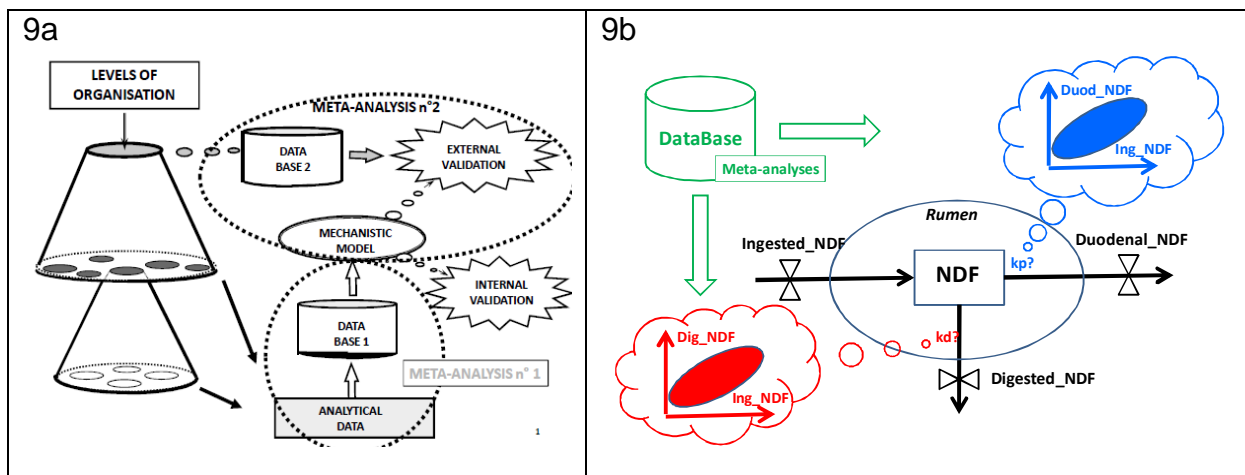
Figure 8: Meta-analysis of spatio-temporal scaling of Dry Matter Intake (DMI) and Time in grazing ruminants



Meta-analysis and mechanistic modelling

One of the main interest of meta-analysis is found in its possibility of combination with mechanistic modelling as Sauvant and Martin (2006) have pointed out. In this publication it was suggested to apply meta-analysis at the two different phases of construction and evaluation of the developed mechanistic model (**Figure 9a**). First, at the underlying level meta-analyses of a first database allows to obtain mean and realistic values for the key parameters and relationships used to build the mechanistic model. Secondly, at the most integrated level, meta-analyses of a second database obtained at this level where the system is only a subpart or an element allows to assess the global validity of the model.

Figure 9: Articulations between mechanistic modelling and meta-analysis(9a) and principle of mechanistic modelling of digestion of ruminal NDF based on structural equations (9b).



Since then, other synergies have been proposed between mechanistic modelling and meta-analytic approaches. Thus, it is possible to build mechanistic models using a top-down approach by relying on intra-experiment regressions $Y_i = f(X_i)$ used as "structural equations" to adjust values of the underlying parameters related to flows and compartments (**Figure 9b**). These structural regressions are established from meta-analyses of databases of flows of inputs and outputs. Such an approach has already been used to construct a simple mechanistic models of fiber digestion in the rumen (6 compartments) associated with meta-analyses of a database of $n_{exp}=219$ and $n=552$ (Sauvant and Mertens, 2009). A more recent and complicated example concerned a mechanistic model of the liver (Bahloul, 2014 and submitted). In this context, the domain of validity of such models is directly linked to the range of data included in the meta-analysis, and thus, used for the calibration of the structural equations, integrated into the mechanistic model. In the first example of fiber digestion, the levels of dry-matter intake (expressed as percentage of BW) determines this domain of validity whereas in the second example (liver model), this domain is defined by the range of major splanchnic flows of nutrients.

Another example is the modeling of dietary calcium and phosphorus fate in growing pigs (Létourneau-Montminy et al., 2015), where meta-analysis has been used to parameterized phosphorus and calcium digestive absorption (Létourneau-Montminy et al., 2012) and then to validate the model.

The recent INRA 2018 feed unit system is another example of synergy between meta-analyses and mechanistic modelling. Indeed, to update the system specific regressions were obtained from meta-analyses of sub-databases focused on various feeding factors: for instance a database with experiments focused on the effect of various nutrients (starch, protein, lipids) to build regressions of those respective nutrient digestive flows (see example of starch mentioned above). The consistency of these equations was checked on a digestive mechanistic model that was calibrated on structural equations of fluxes derived from meta-analysis of the literature (Sauvant and Nozière, 2012, 2016). Moreover if the data and meta-analysis are representative of the feeding practices encountered in field conditions, this model is capable of application to a wide range of practical situations. In addition, it makes it possible to tackle issues linked with efficiency and robustness (Sauvant 2019).

The futures of meta-analysis

The future of meta-analysis will depend on their ability to generate progress in evolving areas of animal sciences. One can think of the Livestock Farming System domain because the data collected in the field are likely to be rather heterogeneous because they come from many different sources and time frequencies (Gonzalez et al., 2018). In this area an interest of meta-analyses could be, thanks to the proposed empirical models, to diagnose in real-time the measurements. In addition, it is likely that meta-analysis will highlight some key parameters that could stimulate technology to progress at capturing these parameters.

Another rapidly evolving sector concerns all genome-focused studies, so far the information is sufficiently sectorized, homogeneous and standardized to be processed by multivariate statistical analysis methods. It is possible that the search for ever greater integration of data will require specific processes of data-fusion using methods already applied in the context of meta-analysis (Zhao et al., 2019).

Another phenomenon will increase the interest of the meta-analysis in the future. Indeed certain experimental techniques are considered too invasive for the animals and will most likely not be authorized to be practiced any more (i.e. digestive cannulas ...). As a result, the data published before with these techniques have a priori an increased scientific value and the challenge is to be able to interpret them with great completeness by meta-analysis.

In the longer term, one of the challenges of meta-analysis will be related to the possibility of automating some of the key steps through the use of artificial intelligence methods.

Conclusions

The results of numerous publications incorporating meta-analysis approaches show that these methods are now widely accepted and applied in animal sciences, especially in nutrition. The methods applied remain largely the same, however the required levels of reporting, traceability, statistic increases and this will ensure more transparency in the process and a better repeatability between analysts. In particular this requirement relates to the choice of the selected publications, as this defines the degree of representativeness of the work with respect to the applications, to the construction and the coding of the database, and to the study of the meta-design. It also relates to the systematic analysis of interfering factors in situations where many candidate independent variables are available.

The debate about the choice between fixed and random effects remains, however our comparisons show that in the vast majority of situations the adjusted values are the same and ultimate conclusions are not influenced by this choice. In concrete terms it is mainly the partition of variance across the effects which seems to be altered by the choice. In particular, we have shown that treatments and experiments that may be identified as outliers are the same independently of the choice between fixed or random. One promising aspect of the development of meta-analyses is related to their implications in other modern approaches such as help in systemic approaches and mechanistic modeling as well. One of the future challenge for meta-analyses will be to demonstrate their interest in the development of precision livestock farming as well as in the processing of large data sets when they faced to heterogeneity and lack of data linked to variations in measurement methods. Otherwise, until now, meta-analysis has mainly been applied to treatments that represent averages of several individuals. It seems that, given growing the interest of phenotype studies, it would be desirable to be able to assess the advantages and limits of meta-analyses to interpret individual laboratory databases by grouping the results of various experiments in which the measured characters are not systematically the same.

References

- Agastin A, Sauvant D, Naves M, and Boval M 2014. Influence of trough versus pasture feeding on average daily gain and carcass characteristics in ruminants: A meta-analysis. *Journal of Animal Science* 92, 1173-1183.
- Bahloul L 2014. Modélisation mécaniste de la émission splanchnique de nutriments énergétiques chez le ruminants. Thèse AgroParisTech. 335 pages
- Bahloul L, Sauvant D, Loncke C, Chartoire M, Vernet J, Lapierre H, Nozière P and Ortigues-Marty I 2019. A novel approach combining meta-analysis with

- mechanistic modeling to predict hepatic nutrient fluxes in ruminants. (Submitted in Journal of Animal Science).
- Berlin J.A., 1995, Benefits of Heterogeneity in Meta-analysis of Data from Epidemiologic Studies. *Am J Epidemiol* 142, No. 4, 384-387.
- Borenstein M, Hedges L and Rothstein H 2007. Meta-Analysis Fixed effect vs. random effects. www.Meta-Analysis.com. 159 p.
- Borenstein M, Hedges LV, Higgins JPT and Rothstein HR 2009. Meta-Regression, Chapter 20 (17p), in "Introduction to Meta-Analysis". Ed Wiley & Sons, Ltd.
- Boval M and Sauvant D 2019. Ingestive behaviour of grazing ruminants: meta-analysis of the components of bite mass. *Animal Feed Science and Technology* 251, 96-111.
- Boval M, Edouard N and Sauvant D 2015. A meta-analysis of nutrient intake, feed efficiency and performance in cattle grazing on tropical grasslands. *Animal* 9, 973-982.
- D'Alexis S, Sauvant D and Boval M 2014. Mixed grazing systems of sheep and cattle to improve liveweight gain: a quantitative review. *Journal of Agricultural Science* 152, 655-666.
- Daniel JB, Friggens NC, Chapoutot P, Van Laar H and Sauvant D 2016. Milk yield and milk composition responses to change in predicted net energy and metabolizable protein: a meta-analysis. *Animal* 10, 1975-1985.
- Daniel JB, Friggens NC, Van Laar H and Sauvant D 2017. Milk protein yield response to change in predicted net energy and metabolizable protein supply: Influence of dry-matter intake response. Annual Meeting of the European Association for Animal Production (EAAP), Tallinn, EST.
- Dragomir C, Sauvant D, Peyraud JL, Giger-Reverdin S, Michalet-Doreau B 2008. Meta- Analysis of 0-8 hours post-prandial kinetics of ruminal pH. *Animal*. 2008 2: 1437-48.
- Eugene M, Archimede H and Sauvant D 2004. Quantitative meta-analysis on the effects of defaunation of the rumen on growth, intake and digestion in ruminants. *Livestock Production Science* 85, 81-97.
- Faverdin P, Sauvant D, Delaby L, Lemosquet S, Daniel JB and Schmidely P 2018. Dry matter intake and milk yield responses to dietary changes. Chap 9. In "INRA Feeding System for Ruminants", 149-176. WAP, NLD.
- Glass GV 1976. Primary, secondary and meta-analysis of research. *Education Research* 5, 3. 8.
- Gonzalez LA, Kyriazakis I and Tedeschi LO 2018. Review: Precision nutrition of ruminants: approaches, challenges and potential gains. *Animal* 12, S246-S261.
- Greenland S., 1987, Quantitative methods in the review of epidemiologic literature. *Epid. Rev.*, Volume 9,1,1. 30.
- Higgins JP, Thompson SG, Deeks JJ and Altman DG, 2003. Measuring inconsistency in meta-analyses. *Brit.Med.J.*;327:557. 60.
- Hillebrand H 2009. Meta analysis of grazer control of periphyton biomass across aquatic ecosystems. *J. of Phycology*, 45 (4), 798-806.
- INRA, 2018,"INRA Feeding System for Ruminants, 643p, Wageningen Academic Publishers, Wageningen, NLD.

- Letourneau-Montminy MP, Cirot O and Lambert W 2018. Effect of low protein diets on nitrogen utilization, daily water consumption, and litter quality in broilers through meta-analysis approach. PSA meeting symposium-Aminoacids and low protein diets: Benefits for performance, meat quality, environment, health and welfare of poultry birds, Poultry Science, Vol. 97, E-Suppl. 1
- Letourneau-Montminy MP, Jondreville C, Sauvant D and Narcy A 2012. Meta-analysis of phosphorus utilization by growing pigs: effect of dietary phosphorus, calcium and exogenous phytase. *Animal* 6, 1590-1600.
- Letourneau-Montminy MP, Narcy A, Dourmad JY, Crenshaw T D and Pomar C 2015. Modeling the metabolic fate of dietary phosphorus and calcium and the dynamics of body ash content in growing pigs. *Journal of Animal Science* 93, 1200-1217.
- Letourneau-Montminy MP, Narcy A, Lescoat P, Bernier JF, Magnin M, Pomar C, Nys Y, Sauvant D and Jondreville C 2010. Meta-analysis of phosphorus utilisation by broilers receiving corn-soyabean meal diets: influence of dietary calcium and microbial phytase. *Animal* 4, 1844-1853.
- Letourneau-Montminy MP, Narcy A, Lescoat P, Bernier JF, Magnin M, Sauvant D, Jondreville C, and Pomar C 2011. Modeling the fate of dietary phosphorus in the digestive tract of growing pigs. *Journal of Animal Science*, 89, 3596. 3611
- Littell RC, Henry PR, and Ammerman CB 1998. Statistical Analysis of Repeated Measures Data Using SAS Procedures. *J. Anim. Sci.* 76:1216. 1231.
- Littell RC, Milliken GA, Stroup WW and Wolfinger R.D 1996. SAS, a System for Mixed Models. pp 31-63. SAS Inst. Inc., Cary, NC.
- Loncke C, Nozière P, Bahloul L, Vernet J, Lapierre H, Sauvant D and Ortigues-Marty I. 2015. Empirical prediction of net splanchnic release of ketogenic nutrients, acetate, butyrate and β -hydroxybutyrate in ruminants: A meta-analysis. *Animal* 9, 449-463.
- Loncke C, Ortigues-Marty I, Vernet J, Lapierre H, Sauvant D and Nozière P 2009. Empirical prediction of net portal appearance of volatile fatty acids, glucose, and their secondary metabolites (β -hydroxybutyrate, lactate) from dietary characteristics in ruminants: a meta-analysis approach. *Journal of Animal Science* 87, 253. 268.
- Martin O and Sauvant D 2002. Meta-analysis of Input/Output Kinetics in Lactating Dairy Cows. *Journal of Dairy Science*, 85, 3363. 3381.
- Offner A and Sauvant D 2004. Prediction of in vivo starch digestion in cattle from in situ data. *Animal Feed Science and Technology* 111, 41-56.
- Pearson K 1904. Report on certain enteric fever inoculation statistics. *Brit.Med.J.*, 3, 1243-1246.
- Petitti DB 2001. Approaches to heterogeneity in meta-analysis. *Statist. Med.*, 20, 3625. 3633
- Philibert A, Loyce C and Makowski D 2012. Assessment of the quality of meta-analysis in agronomy. *Agri., Ecosyst. and Env.* 148, 72. 82
- Phillips CJC 2005. Meta-analysis-a systematic and quantitative review of animal experiments to maximise the information derived. *Anim.Welfare*, 14, 333-338.
- Salah N, Sauvant D and Archimede H 2014. Nutritional requirements of sheep, goats and cattle in warm climates: a meta-analysis. *Animal* 8, 1439-1447.
- Salah N, Sauvant D and Archimede H 2015. Response of growing ruminants to diet in warm climates: a meta-analysis. *Animal* 9, 822-830.

- Sauvant D 2019. Modelling efficiency and robustness in ruminants, the nutritional point of view. *Animal Frontiers*. 9. 2, 60-67.
- Sauvant D and Martin O 2004. Empirical modelling through meta-analysis vs mechanistic modelling. In *Nutrient digestion and utilization in farm animals: modelling approaches* (ed. E Kebreab). CAB International, London.
- Sauvant D and Nozière P 2012. Modèle intégratif du tube digestif intégrant les interactions digestives, les flux de nutriments d'intérêt et compatible avec les systèmes UF et PDI. *Rencontres Recherches Ruminants* 19, 181-184.
- Sauvant D and Nozière P 2016. Quantification of the main digestive processes in ruminants: the equations involved in the renewed energy and protein feed evaluation systems. *Animal* 10, 755-770.
- Sauvant D and Yang WZ 2011. Meta-analysis of the additivity between two dietary fibres in dairy cows (french). *Rencontres Recherche Ruminants*, 18, 125.
- Sauvant D, Giger-Reverdin S, Serment A and Broudiscou L 2011. Influences of diet and rumen fermentation on methane production by ruminants. *Inra Productions Animales* 24, 433-446.
- Sauvant D, Nozière P and Ortigues-Marty I 2018. Chapter 6. Energy expenditures, efficiencies and requirements. In *INRA Feeding System for Ruminants*, (ed. Inra), pp. 91-118, Wageningen Academic Publishers, Wageningen, NLD.
- Sauvant D, Schmidely P, Daudin J J and St-Pierre NR 2008. Meta-analyses of experimental data in animal nutrition. *Animal* 2, 1203-1214.
- Srednicka-Tober D, Baranski M, Seal C, Sanderson R, Benbrook C, Steinshamn H, Gromadzka-Ostrowska J, Rembalkowska E, Skwarlo-Sonta K, Eyre M, Cozzi G, Larsen MK, Jordon T, Niggli U, Sakowski T, Calder PC, Burdige GC, Sotiraki S, Stefanakis A, Yolcu H, Stergiadis S, Chatzidimitriou E, Butler G, Stewart G and Leifert C 2016. Composition differences between organic and conventional meat: a systematic literature review and meta-analysis. *British Journal of Nutrition* 115, 994-1011.
- St-Pierre NR 2001. Invited review: Integrating quantitative findings from multiple studies using mixed model methodology. *Journal of Dairy Science* 84, 741-755.
- Sutton AJ and Higgins JP 2008. Recent developments in meta-analysis. *Statist.Med.* 27, 625-650.
- Thompson SG 1994. Systematic review - why sources of heterogeneity in metaanalysis should be investigated. *British Medical Journal* 309, 1351-1355.
- vanHouwelingen H C, Arends LR and Stijnen T 2002. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 21, 589-624.
- Vernet J and Ortigues-Marty I 2006. Conception and development of a bibliographic database of blood nutrient fluxes across organs and tissues in ruminants: data gathering and management prior to meta-analysis. *Reproduction Nutrition Development* 46, 527-546.
- Yates F and Cochran WG 1938. The analysis of groups of experiments. *J.Agric.Sci.*, 28, 556-580.
- Zhao J, Sauvage C.1,6, Zhao J., Bitton F., Bauchet G.1,8, Liu D., Huang S., Tieman D.M., Klee H.J. & Causse M., 2019, Meta-analysis of genome-wide association

studies provides insights into genetic control of tomato flavor. Nature Communications, 1-12, | <https://doi.org/10.1038/s41467-019-09462-w>

Zouaoui M, Lambert W, Létourneau-Montminy MP 2019. Meta-analysis of the response of broilers to dietary valine: impact of other branched chain amino acids, PSA Annual Meeting, P. Sci Vol. 98, E-Suppl. 1, In press