



This is an original version of a chapter published in  
Proceedings of the 5th Nordic Feed Science Conference.

Citation for the published publication:

Eriksson, Torsten. (2014) Simple Excel VBA application for Monte Carlo simulation with ad hoc made spreadsheet models. In: Udén, Peter; Eriksson, Torsten; Rustas, Bengt-Ove; Müller, Cecilia; Spröndly, Rolf; Thomas, Pauly; Emanuelson, Margareta; Holtenius, Kjell; Patel, Mikaela; Åkerling, Maria and Jansson, Anna (eds) *Proceedings of the 5th Nordic Feed Science Conference*. Uppsala: Department of Animal Nutrition and Management, Swedish University of Agricultural Sciences, pp 138-142.

Epsilon Open Archive <http://epsilon.slu.se>



# Proceedings of the 5<sup>th</sup> Nordic Feed Science Conference, Uppsala, Sweden



**Nordic Feed Science Conference**

**10 - 11 June**

**2014**

[www.slu.se/nordicfeedscienceconference](http://www.slu.se/nordicfeedscienceconference)

---

**Institutionen för husdjurens  
utfodring och vård**

**Swedish University of Agricultural Sciences  
Department of Animal Nutrition and Management**

**Rapport 290  
Report**

**Uppsala 2014**

ISSN 0347-9838  
ISRN SLU-HUV-R-290-SE

---

## **Simple Excel VBA application for Monte Carlo simulation with ad hoc made spreadsheet models**

T. Eriksson

*Department of Animal Nutrition and Management, Swedish University Agricultural Sciences, Kungsängen Research Center, SE-753 23 Uppsala, Sweden.*

*Correspondence: torsten.eriksson@slu.se*

### **Introduction**

Researchers in animal science as well as in many other disciplines use spreadsheet models in a multitude of ways. There are large, well organized models created with considerable efforts. But, the majority of models are ad hoc made applications quickly put together when checking up some idea that struck the researcher examining the outcome of an experiment or reviewing a manuscript. These models may use only a handful of cells, but sometimes evolve into huge poorly organized “spaghetti models” as more factors are taken into account.

A typical issue when examining complex relationships involving multiple factors is what effect they may have on the variation of the response variables. This could be tested by Monte Carlo simulation, where, for instance, multiple model parameters are allowed to vary according to a chosen distribution. The outcome of thousands of simulations generates a distribution in response variables which could be regarded as an extended sensitivity analysis. This paper describes a simple spreadsheet application in Microsoft Excel that easily can be connected to existing ad hoc spreadsheet models for generating Monte Carlo simulations. The application is free for download and use.

### **The application**

The application (APP) was made in Excel 2010 and consists of a worksheet and a short VBA code (macro) operated by a button on the worksheet. The VBA code allows simulations to be run automatically any chosen number of times (thousands). The actual value of each parameter in each simulation iteration as well as the corresponding values of response variables are logged and forms the outdataset from the simulation. The worksheet could either be inserted into the Excel file containing the existing model or, the model could be copied into the file with the APP. In both cases, the file must be saved as a macro enabled workbook.

The APP was made as transparent as possible, to facilitate understanding of the construction and make it easier for advanced users to modify it. The present worksheet has seven rows with constants or formulae (Figure 1), of which three require user activity for each new simulation setup. The other rows could be changed by expanding or diminishing the range of cells in use. The worksheet in the example is set up for a maximum of 30 model parameters and storage of these and a maximum of 20 model output variables.

Random coefficients are generated in Rows 1 and 2. Row 1 has a range of cells holding constants. In the example, there are 800 cells with values following a normal distribution with mean = 0 and  $s = 1$ , i. e., the values vary from approx. -3 to 3. The values were generated by Excel Analysis ToolPak but could also be imported from other sources or replaced by data following some other distribution if considered more suitable. Row 2 contains RANDBETWEEN and OFFSET formulas to retrieve a value from a randomly chosen cell in the range of normally distributed constants in Row 1. The value is replaced at each

recalculation of the worksheet, i.e. at any manual change to the worksheet, when F9 key is pressed or by 'Calculate' in a macro.

Row 3 is for user input of starting values for model parameters to be varied in the simulations. In the example, the average value for each parameter is used. The standard deviation for each parameter to be varied is entered in Row 4.

In Row 5, the random coefficients generated in the first step are multiplied with user entered standard deviations. This gives the absolute deviation for each parameter in the current simulation. In Row 6, all values are checked for non-negativity and added to the starting values to yield parameter values to be used in the simulation. If negative, a zero value is used.

Row 7 contains both current values of the parameters and model variable output data of interest, i.e. result of the simulation. Variable values are linked with Row 6 and with the existing model. Output data from the existing model are linked to cells in Row 7 to the right of the parameters. The row is logged below in a designated cell range after each simulation.

	Random coefficient generation		
Row 1	Cells populated with standardized distribution (In example normal distribution with $x = 0$ ; $s = 1$ )	-2.7107 .....	"
Row 2	Coefficients retrieved from randomly chosen cells containing the distribution	-0.273 .....	"
	Input of values and known/assumed variation for model parameters of interest		
	Labels for model parameters	Ash, g/kg DM .....	"
Row 3	Original values for model parameters that should vary in the simulation	22.9 .....	"
Row 4	Known/assumed standard deviation for each varying model parameter	2.0	
	Calculation of parameter value to be used in the simulation		
Row 5	Random coefficients are multiplied with the known/assumed parameter standard deviations	-0.5 .....	"
Row 6	Results are added to the original values and checked for non-negativity	22.4 .....	"
	Input in model of parameters with simulated variation and retrieval of results		
	Parameter values to be used are linked from Row 6 and further to the formulas in the model. Outdata cells of interest from the model are linked to the same row.		
Row 7	The row is after each recalculation copied to a database.	22.4 .....	"

**Figure 1** Structure of the spreadsheet for generating Monte Carlo simulations from existing Excel models. Data continue to the right in as many columns as needed. User activity is required for entering parameter values and standard deviations in Rows 3 and 4, respectively and for linking cells to Row 7. Row 7 is logged after each simulation.

### An example

The example chosen is the summative effect of intra-laboratory analytical variation of laboratory estimated metabolizable energy (ME) content of feeds and resulting ME intake of a dairy cow diet. Proximate analysis of concentrates with a set of tabulated coefficients for digestibility and metabolizability has been used in Sweden since 1967. Together with forage ME estimation from in vitro organic matter digestibility, it has formed the Swedish system for ruminant ME calculation from the late 1970's and onwards and has only recently been partly replaced by the Norfor model.

A diet with barley, rapeseed cake and grass silage was used. Initially, variation in the ME concentration of barley was investigated and in the next, total ME intake from the diet and the ME allowable production of energy corrected milk (ECM) according to Spörndly (2003)

was simulated. Table 1 shows analytical values together with calculated ME concentration of the feeds and standard deviation for repeated analyses of control samples at the Kungsängen Research Centre laboratory (Verner et al., 2012).

**Table 2** Metabolizable energy calculation from proximate analysis of concentrates and in vitro organic matter digestibility of forage for the feeds used in the simulation examples. Intra-laboratory standard deviations (reproducibility over time for the Kungsängen Research Centre laboratory) that were used in the simulations are included

Item	Barley	Rapeseed cake	Grass silage	Reproducibility (SD between repeated analyses)
<i>Proximate analysis</i>				
Ash, g/kg DM	28	66	60	2
Crude protein, g/kg DM	122	339		0.75
Crude fat, g/kg DM	27	169		1.10
Crude fiber, g/kg DM	62	137		4.0
Nitrogen free extractives, g/kg DM <sup>a</sup>	761	289		
<i>Forage OMD and ME calculation</i>				
In vitro OMD of forage, g/kg OM			873	9.2
MJ ME/ kg DM	13.16	15.46	11.34	

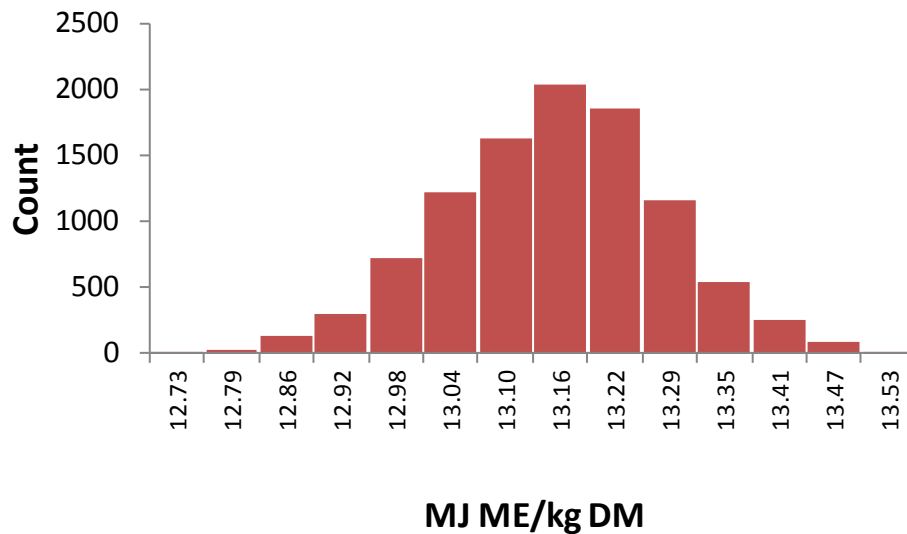
<sup>a</sup>Calculated as a residual: 1000-ash-crude protein-crude fiber

**Table 2** Total ME intake and production potential of energy corrected milk (ECM) for the diet in the simulation example

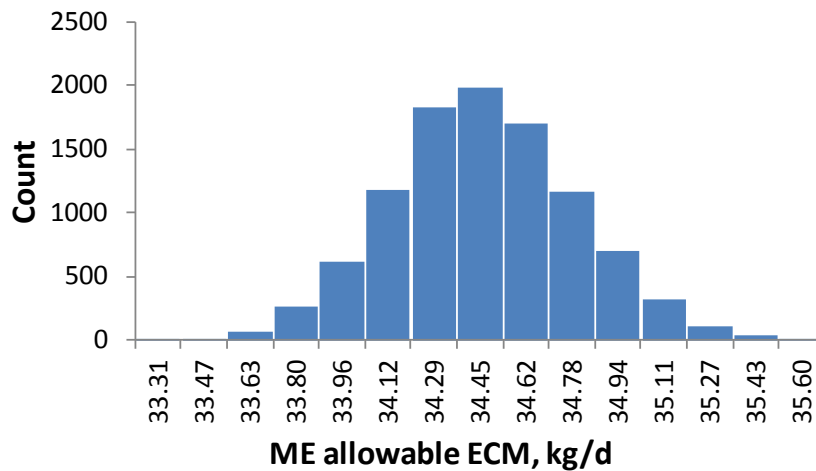
Item	Barley	Rapeseed cake	Grass silage	Total diet
MJ ME/ kg DM	13.16	15.46	11.34	
Intake, kg DM/d	6	2	12	20
Total intake, MJ ME/d				246
ME allowable ECM, kg/d				34.5

## Results

Simulations were run on a standard laptop computer with 2.1 GHz processor and 8 GB RAM memory, which required 150 seconds for 10000 simulations of barley ME concentration with four parameters. For the total diet, ten parameters were used and the simulation required 165 seconds for 10000 simulations. Correlations among parameters were in absolute values < 0.03. The range of ME values for barley was about 0.8 MJ/kg DM (Figure 2). The total diet ranged from 239 to 252 MJ ME/d with a standard deviation of 1.8 MJ. The estimated resulting ECM production supplied (Figure 3) varied over a range of 2.35 kg/d as an effect of laboratory analytical variation.



**Figure 2** Metabolizable energy (ME) concentration from 10000 simulations with the example barley. All variation is for the result of known variation in laboratory analytical data. Range 12.72 – 13.54, s 0.123. Class width is 0.5 s.



**Figure 3** Energy corrected milk (ECM) production as allowed by metabolizable energy (ME) intake from 10000 simulations with the example diet. All variation is due to the known variation for laboratory analyses of the three feeds in the diet. Range 33.29 – 35.64, s 0.327. Class width is 0.5 s.

### Discussion

The simulations displayed the possible variation in outcome when four or ten parameters are allowed to vary independently following normal distribution. It would have been possible to let a much larger number of parameters vary if feasible. In this specific case, the assumption of no covariance between variables is probably correct, because it deals with random analytical variation at a laboratory. However, when applying this approach to other models,

the possibility of covariance between variables should always be taken into account. Consider for instance a ruminant nutrition model, where there most likely exists covariance between ruminal digestibility and small intestinal digestibility of a feed component.

The random generator in Excel has occasionally been criticized (McCullough, 2008), but the absence of correlation between variables does not suggest systematic errors. It would be possible to extend the population of coefficients for sampling and also to perform repeated recalculations in each simulation step, although simulation time would increase.

The situation simulated here has likely occurred in many dairy cow experiments and would have had minor consequences for result interpretation if: 1) the experiment was of a change-over design and random analytical differences would just add to period effects or 2) feed samples from the entire experiment were analysed in the same batch. However, if none of these conditions are fulfilled, the repeatability variation at a research laboratory could cause differences of the magnitude, shown above.

Analytical variation of a control sample will not reflect what the variation will be in all sample types and it is likely that variation actually is larger in some cases. Reproducibility among research laboratories or commercial laboratories is generally greater and if such data is mixed within an experiment, variation would most likely increase considerably.

### **Conclusion**

A relatively simple Excel APP can be used for investigating the outcome when multiple parameters are varied independently in simulations. They can easily be linked to existing Excel models by the user.

### **References**

McCullough, B.D. 2008. Microsoft Excel's 'Not The Wichmann-Hill' random number generators. *Comp. Stat. Data Anal.* 52, 4587–4593.

Spörndly, R. 2003.(Ed) *Fodertabeller för idisslare* (in Swedish). Department of animal nutrition and management. Swedish University of Agricultural Sciences. SLU/HUV Report 257, pp, 7-15.

Verner, E., Ericson, B. & Udén, P., 2012. The variation of control samples in the analyses of crude protein, neutral detergent fibre, in vitro organic matter digestibility and crude fat. 3<sup>rd</sup> Nordic Feed Science Conference, Uppsala, 28-29 of June 2012. SLU/HUV Report 280, p 27-30.

The Excel APP is downloadable at: [www.slu.se/en/torsten-eriksson](http://www.slu.se/en/torsten-eriksson)