



Methods for simulating nutritional requirement and response studies with all organisms to increase research efficiency

Dmitry Vedenov¹, Rashed A. Alhotan², Runlian Wang^{2,3} and Gene M. Pesti^{2*}

¹Department of Agricultural Economics, Texas A&M University, College Station, TX 77843, USA

²Department of Poultry Science, The University of Georgia, Athens, GA 30602, USA

³Department of Animal Science, Guangdong Ocean University, Zhanjiang 524088, People's Republic of China

(Submitted 21 July 2016 – Final revision received 11 December 2016 – Accepted 23 December 2016 – First published online 27 February 2017)

Abstract

Nutritional requirements and responses of all organisms are estimated using various models representing the response to different dietary levels of the nutrient in question. To help nutritionists design experiments for estimating responses and requirements, we developed a simulation workbook using Microsoft Excel. The objective of the present study was to demonstrate the influence of different numbers of nutrient levels, ranges of nutrient levels and replications per nutrient level on the estimates of requirements based on common nutritional response models. The user provides estimates of the shape of the response curve, requirements and other parameters and observation to observation variation. The Excel workbook then produces 1–1000 randomly simulated responses based on the given response curve and estimates the standard errors of the requirement (and other parameters) from different models as an indication of the expected power of the experiment. Interpretations are based on the assumption that the smaller the standard error of the requirement, the more powerful the experiment. The user can see the potential effects of using one or more subjects, different nutrient levels, etc., on the expected outcome of future experiments. From a theoretical perspective, each organism should have some enzyme-catalysed reaction whose rate is limited by the availability of some limiting nutrient. The response to the limiting nutrient should therefore be similar to enzyme kinetics. In conclusion, the workbook eliminates some of the guesswork involved in designing experiments and determining the minimum number of subjects needed to achieve desired outcomes.

Key words: Simulation: Nutritional requirements: Excel: Models

The concept of nutrient requirements is central to the field of nutrition. Nutritional response kinetics refers to the basic shape of the curve describing the relationship between nutrient input levels (concentration or intake) and responses (growth, efficiency, metabolism, lactation, egg production, etc.). The particular choice of the mathematical response model to describe nutritional kinetics is, therefore, very important to nutritionists who need to estimate requirements of any organism. While some nutritionists consider that there should be a smooth transition from the ascending to the plateau portions of the response^(1–4), the practical approaches are often to use multiple range tests, polynomials or the 'broken-line' spline model, to estimate the requirements for maximum performance, efficiency of nutrient utilisation, etc.^(4–8). Just a very few examples of response models that have been described in this manner are weight gain *v.* nitrogen intake and isoleucine retention *v.* dietary casein level in rats, bone ash *v.* vitamin D level and blood

clotting time *v.* vitamin K in chicks⁽¹⁾, whole body protein *v.* sulphur-containing amino acids in pigs⁽⁵⁾, cataracts *v.* histidine levels⁽⁶⁾ and energy metabolism *v.* arginine⁽⁷⁾ in salmon diets, and the protein requirements of healthy pregnant women⁽⁸⁾. In human nutrition, the practice has been to relate requirements or recommended allowances to clinical status, growth and body fluid composition. In agricultural animals and fish, a balance must be struck between feed costs and time to reach market weights and carcass composition⁽¹⁾. In crop production, a balance can be found using kinetic response models to maximise the difference between crop value and fertiliser costs.

To help nutritionists design more efficient nutritional experiments, we developed an Excel-based simulation program. The program, called the Nutritional Response Determination Optimization workbook (NuRDO.xls), simulates a series of experimental responses around a pre-selected 'true' model, which can be then used to estimate parameters and

Abbreviations: BLL, broken-line spline model with linear ascending segment or broken-line linear model; BLQ, broken-line spline model with quadratic ascending segment or broken-line quadratic model; NuRDO.xls, Nutritional Response Determination Optimization workbook; QP, quadratic regression; RC, rate constant; SK, saturation kinetics model.

* **Corresponding author:** G. M. Pesti, email gpesti@uga.edu

requirements. In particular, NuRDO can be used to simulate responses based on quadratic polynomials (QP, $y = b_0 + b_1x + b_2x^2$), broken-line spline models with either linear (BLL, $y = \text{maximum}$ if $x > \text{requirement}$, $\text{maximum} + \text{rate constant} \times (\text{requirement} - x)$ if $x \leq \text{requirement}$) or quadratic (BLQ, $y = \text{maximum}$ if $x > \text{requirement}$, $\text{maximum} + \text{rate constant} \times (\text{requirement} - x)^2$ if $x \leq \text{requirement}$) ascending segments, or the saturation kinetics (SK, $y = (\text{intercept} \times \text{rate constant} + \text{maximum} \times x^{\text{kinetic order}}) / (\text{rate constant} + x^{\text{kinetic order}})$) model⁽¹⁾ with parameters input by the user (Table 1; Fig. 1). For all models, $y = \text{response variable}$, $x = \text{nutrient level}$, $b_0 = \text{intercept}$, b_1 and $b_2 = \text{regression coefficients}$, while other variables are parameters.

NuRDO can be used to evaluate the effects of experimental design choices on the expected confidence in requirements and/or other parameter estimates. The objectives of the studies reported here were to determine how standard errors of the estimates of nutritional requirements are affected by (1) the number of simulated experiments; (2) the number of nutrient levels and replicates (reps) per level; (3) the range of nutrient levels above and below the actual requirement; and (4) the variability of simulated responses (CV) and the shape of the 'true' response model (rate constant or RC).

The workbook should be most helpful in planning experiments to maximise the efficiency of resource use and to justify the number of experimental subjects required to achieve desired confidence in conclusions from experiments.

Methods

NuRDO is implemented as a Microsoft Excel workbook that allows the user to design a nutritional requirement experiment and simulate its outcomes (<http://www.poultry.uga.edu/extension/PoultryNutrition.htm>). First, the user selects the desired number

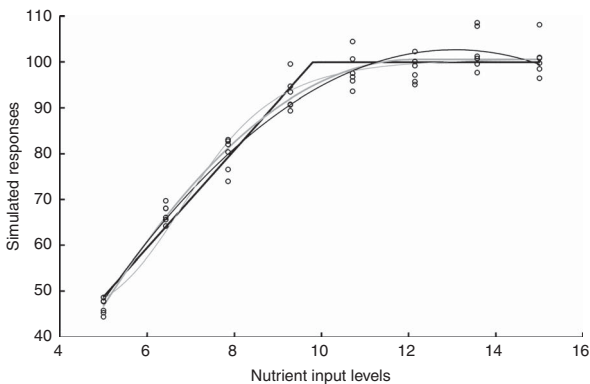


Fig. 1. Example of Nutritional Response Determination Optimization (NuRDO) simulation of data from an experiment with eight nutrient levels between 5 and 15 with a maximum response of 100 and a rate constant of -10 . The models fitted are the quadratic polynomial (QP, $y = b_0 + b_1x + b_2x^2$), broken-line spline models with either linear (BLL, $y = \text{maximum}$ if $x > \text{requirement}$, $\text{maximum} + \text{rate constant} \times (\text{requirement} - x)$ if $x \leq \text{requirement}$) or quadratic (BLQ, $y = \text{maximum}$ if $x > \text{requirement}$, $\text{maximum} + \text{rate constant} \times (\text{requirement} - x)^2$ if $x \leq \text{requirement}$) ascending segments and the saturation kinetics (SK ($\text{intercept} \times \text{rate constant} + \text{maximum} \times x^{\text{kinetic order}} / (\text{rate constant} + x^{\text{kinetic order}})$). For all models, $y = \text{response variable}$, $x = \text{nutrient level}$, $b_0 = \text{intercept}$, b_1 and $b_2 = \text{regression coefficients}$, other variables are parameters. ○, Simulation; —, BLL; - - -, BLQ; ···, QP; - · - ·, SK.

of levels, replicates and the range of inputs. The workbook automatically generates a uniformly spaced grid of input levels. However, the generated levels can be adjusted, for example to cluster them around a suspected inflection point of the BLL or BLQ model. The user then selects the 'true' model from which the responses are simulated and calculates the corresponding model parameters. The available options are BLL, BLQ, QP and SK functional forms. Lastly, the user specifies the CV of simulated responses and the number of experiments to simulate. The workbook then generates a series of responses around the true model (with the random errors drawn from the normal distribution), fits up to four selected models (BLL, BLQ, QP and/or SK), and calculates and outputs the estimated parameters along with their standard errors and 95% CI. The outputs are stored in the workbook and can be used for subsequent analysis.

SE were calculated as square roots of the diagonal elements of the matrix: $\hat{\sigma}^2 [Z'(b)Z(b)]^{-1}$, where $Z(b) = \frac{\partial f(\mathbf{x}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \Big|_b$; $\hat{\sigma}^2 = \frac{[y - f(\mathbf{x}, \boldsymbol{\beta})][y - f(\mathbf{x}, \boldsymbol{\beta})]}{n - k}$, $f(\mathbf{x}, \boldsymbol{\beta})$ is the non-linear function parameters of which was estimated, \mathbf{x} and \mathbf{y} are the inputs and responses from a given experiment, respectively, n is the number of observations and k is the number of parameters. SD were calculated across different simulated experiments, that is: $SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (\beta_i - \bar{\beta})^2}$, where β_i is the estimate of a specific parameter obtained from the simulated experiment i .

Linear regressions were fitted to R^2 's as functions of the total number of pens in the experiment.

Results

As the number of simulated experiments increased, the variation in the mean requirement decreased until thirty simulations were reached (Fig. 2). For the BLL model, increasing the size of the simulated experiment from sixteen to twenty-four pens decreased the SE of the requirement estimate by about 19% (Table 1; 0.46 v. 0.57), from twenty-four to forty-eight pens by about 28% (0.33 v. 0.46) and from forty-eight to ninety-six pens by another 30% (0.23 v. 0.33). SE estimates were much higher

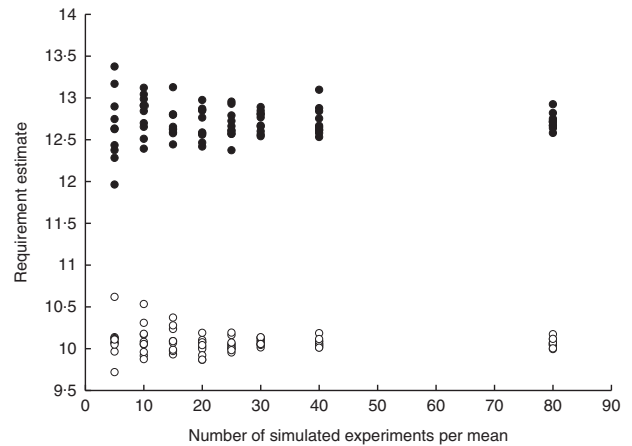


Fig. 2. Influence of number of Nutritional Response Determination Optimization simulations on the precision of requirement estimation. There were ten requirement estimates with five to eighty simulations each. Each point represents one requirement estimate. ●, Broken-line quadratic model; ○, broken-line linear model.

Table 1. Results of several models fitted to data simulated with various combinations of levels and replications per level* (Standard errors and coefficients of determination)

Pens	Levels					Broken-line linear			Broken-line quadratic			QP	SK
	Levels	Reps	Low	High	Break point	Requirement	SE	R ²	Requirement	SE	R ²	R ²	R ²
16	4	4	5	15	10	9.89	0.57	0.91	12.45	1.25	0.91	0.91	0.91
24	6	4	5	15	10	10.09	0.45	0.90	12.82	1.05	0.90	0.90	0.90
24	4	6	5	15	10	10.14	0.48	0.91	12.91	1.12	0.91	0.91	0.91
24	8	3	5	15	10	9.97	0.45	0.89	12.36	0.99	0.88	0.89	0.89
48	6	8	5	15	10	9.92	0.32	0.89	12.50	0.73	0.88	0.88	0.89
48	8	6	5	15	10	10.11	0.34	0.87	12.71	0.77	0.87	0.87	0.87
48	16	3	5	15	10	9.93	0.34	0.85	12.28	0.76	0.84	0.84	0.85
96	8	12	5	15	10	9.97	0.23	0.88	12.43	0.51	0.87	0.87	0.87
96	12	8	5	15	10	9.99	0.23	0.87	12.29	0.51	0.86	0.86	0.86
96	16	6	5	15	10	10.02	0.23	0.86	12.45	0.51	0.86	0.86	0.86
48	6	8	5	15	12	12.08	0.27	0.95	16.17	1.00	0.94	0.94	0.94
48	8	6	5	15	12	12.00	0.28	0.95	15.92	1.02	0.94	0.94	0.94
48	16	3	5	15	12	12.07	0.27	0.94	16.35	1.28	0.93	0.93	0.93

QP, quadratic polynomial; SK, saturation kinetics model; reps, replicates.

* The CV between observations was 8%; equally spaced levels between 5 and 15, for example for four levels = 5.000, 8.333, 11.667, 15.000.

Table 2. Results of several models fitted to data simulated with various ranges of data above and below the simulated nutritional requirement or 'break point'* (Standard errors and coefficients of determination)

Pens	Levels					Broken-line linear			Broken-line quadratic			QP	SK
	Levels	Reps	Low	High	Break point	Requirement	SE	R ²	Requirement	SE	R ²	R ²	R ²
24	4	6	5	15	10	10.19	0.50	0.91	13.20	1.11	0.91	0.91	0.91
24	4	6	6	14	10	10.05	0.51	0.86	12.35	1.08	0.86	0.86	0.86
24	4	6	7	13	10	10.11	FC	0.77	11.91	FC	0.77	0.77	0.77
24	4	6	8	12	10	10.19	FC	0.57	14.10	FC	0.57	0.58	0.57
24	4	6	9	11	10	9.98	FC	0.21	14.60	FC	0.23	0.24	0.23
48	8	6	5	15	10	10.08	0.34	0.87	12.62	0.77	0.86	0.87	0.87
48	8	6	6	14	10	9.87	0.33	0.80	11.80	0.75	0.79	0.79	0.79
48	8	6	7	13	10	10.00	0.34	0.70	11.56	0.84	0.70	0.70	0.69
48	8	6	8	12	10	9.93	0.33	0.55	11.13	0.94	0.54	0.54	0.53
48	8	6	9	11	10	9.94	0.33	0.24	11.36	13.57	0.24	0.23	0.23

QP, quadratic polynomial; SK, saturation kinetics model; reps, replicates; FC, Excel failed to converge on a solution for that data.

* Equally spaced levels between low and high, for example four equally spaced levels between 5 and 15 = 5.000, 8.333, 11.667, 15.000.

(>100%) for the BLQ models. Relative differences for the BLQ model were similar to the BLL model. R² values decreased as the number of pens (total variation) increased from sixteen to forty-eight for all models (BLL, P=0.039; BLQ, P=0.036; QP, P=0.025; SK, P=0.011). If the break-point was not central to the range studied (i.e. break-point of 12, Table 1), the SE decreased for the BLL model, but increased for the BLQ model. R² values increased for all four models when the break point was not in the centre of the range studied. The combination of nutrient levels and number of replications per level made little difference in the precision of the requirement estimates; in this example eight levels of six reps each and sixteen levels of three reps each gave very similar results with SE of 0.27 (to two decimal places).

The influence of range of nutrient levels simulated was dependent on the experimental replication (Table 2). As the range of simulated nutrient levels decreased, R² values decreased and when the range was too small, the models would not converge and thus produced no SE estimates. With forty-eight total observations there was little difference in SE estimates

for the BLL model but for the BLQ model, the estimated SE increased as the simulated interval decreased.

As the CV of the simulated responses increased, the requirement estimates became less precise (Fig. 3 and 4), although there seems to be little difference above CV levels of 6% for either the BLL or BLQ models. The more important factor is the RC, the slope of the ascending portion of the response (Table 3). If the angle between the segments in the BLL model is relatively small, the requirement estimate is more precise. As the angle becomes greater (-20 v. -2), the increase in SE is much greater with higher CV.

Discussion

In planning experiments, it is most helpful for nutritionists to know the expected variation and the shape of the response from historical data; plus they need to have some expectation of the acceptable variation in the requirement estimate. With this basic information, the nutritionist can decide if they have sufficient

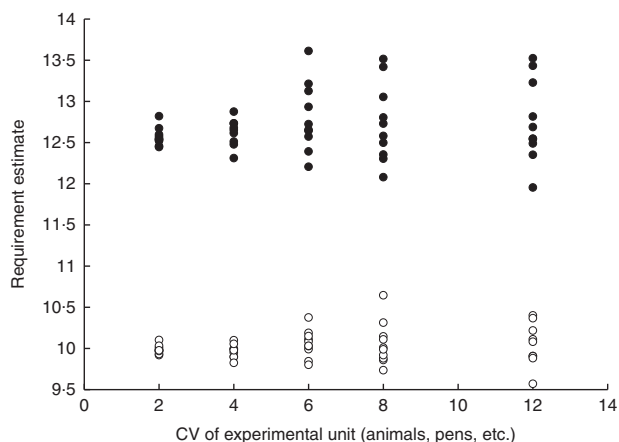


Fig. 3. Effect of the CV of the experimental unit on the precision of requirement estimation. Each point represents one requirement estimate from thirty simulations. ●, Broken-line quadratic model; ○, broken-line linear model.

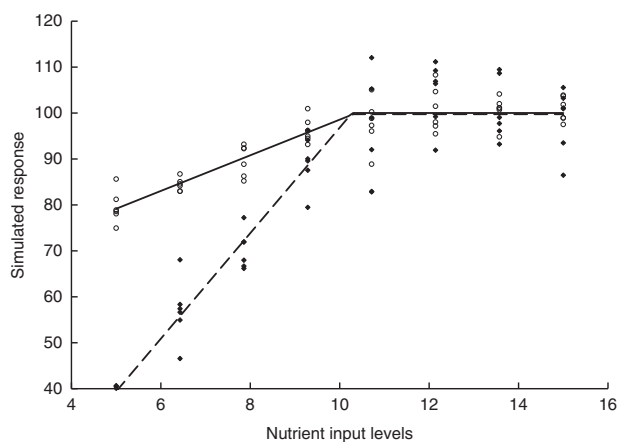


Fig. 4. Example of Nutritional Response Determination Optimization simulations with different rate constants (RC) and different observation to observation CV. ○, Simulations CV = 4%; —, RC = -4; - - -, RC = -12; ◆, simulations CV = 10%.

resources and then how to best allocate the resources to meet their goals. In the past, nutritionists had to make an educated guess on the best way to design their experiments. The NuRDO workbook uses Excel to take at least some of the guesswork out of the process of experimental design. Note that the results from NuRDO simulations will never be exactly the same because they are based on random draws from normal distributions (the same as real experiments have individuals randomly chosen from some greater population). However, with enough simulated experiments, the mean results of multiple simulations should be very similar up to any desired number of decimal places.

The three factors having the largest impact on accuracy of requirement estimates are the CV of the subjects, RC and interval (nutrient levels) to study. If the RC is small and there is a very gradual transition from ascending portion of the response to the plateau portion, it will be difficult to get a precise estimate of the requirement in question regardless of the number of levels and replicates used. This is a result of the requirement estimate being made for the population and not any particular individual(s). The influence of the RC is not just a matter of scaling that can be manipulated in graphing the data. The slope of the regression line is the critical factor and it is represented by the RC in the model used to estimate the SE of the requirement. There is little that can be done to alter the RC. Like the K_m in enzyme kinetics, RC is a property of the system being studied.

On the other hand, the variation between individuals, or the CV, can often be changed to help increase experimental precision. The most common method of decreasing the CV is to choose uniform individuals to study. It is important that conclusions drawn from limited populations are not extrapolated to other populations. Regressions with only individuals near the mean weight, for instance, may not be normally distributed (a transformation may be necessary) and the results may only apply to similar individuals. Choosing one sex or the other, one strain or the other, etc. may be quite useful if it is remembered that the results may only be applicable to the sub-population from which they were derived.

Choosing the appropriate range of nutrient levels to study is an area where NuRDO may be particularly helpful. If there are

Table 3. Results of several models fitted to data simulated with different rate constants and two coefficients of variation* (Standard errors and coefficients of determination)

CV	Levels				Rate constant	Broken-line linear			Broken-line quadratic			QP	SK
	Levels	Reps	Low	High		Requirement	SE	R ²	Requirement	SE	R ²	R ²	R ²
4	8	6	5	15	-2	9.78	0.84	0.52	12.32	1.99	0.52	0.52	0.52
4	8	6	5	15	-4	10.18	0.46	0.80	12.83	1.04	0.79	0.79	0.79
4	8	6	5	15	-6	9.96	0.28	0.91	12.51	0.63	0.90	0.90	0.90
4	8	6	5	15	-8	9.93	0.20	0.95	12.41	0.47	0.94	0.94	0.94
4	8	6	5	15	-10	10.02	0.16	0.97	12.46	0.39	0.96	0.96	0.96
4	8	6	5	15	-20	9.99	0.07	0.99	12.48	0.23	0.99	0.98	0.99
8	8	6	5	15	-2	10.30	1543028	0.22	16.74	57.83	0.21	0.22	0.21
8	8	6	5	15	-4	10.01	0.87	0.51	13.01	2.47	0.50	0.51	0.51
8	8	6	5	15	-6	10.01	0.59	0.70	12.66	1.34	0.69	0.69	0.69
8	8	6	5	15	-8	10.11	0.44	0.81	12.69	0.99	0.80	0.80	0.80
8	8	6	5	15	-10	9.87	0.32	0.87	12.30	0.74	0.86	0.86	0.86
8	8	6	5	15	-20	10.00	0.15	0.97	12.45	0.37	0.96	0.96	0.96

CV = $sd/mean$; QP, quadratic polynomial; SK, saturation kinetics model; reps, replicates.
* Eight equally spaced levels between 5 and 15.

no preliminary or historical data then the range to be studied should be particularly wide to include all possible outcomes. Once there is preliminary or historical data to base experiments on, then NuRDO can be helpful in making sure the range studied is not too narrow. If there is too much concentration near the requirement, then the R^2 may be low and the model may even fail to converge on an SE estimate. While it seems intuitive that replication should be concentrated near the break-point, it is not necessarily helpful if the range studied is narrowed, or if the RC is small.

There are two aspects of NuRDO simulations that are critical for accurate predictions: model fitting and random observation distribution. The model fitting method implemented in Excel by NuRDO gives the same result as SAS^(9,10). Non-linear equations are fitted by iterative methods, trial and error, until a given tolerance is reached to find the coefficients minimising the model sum of squares. While no method is perfect, the method implemented by NuRDO gives the same results as one of the most accepted statistical analytical packages available.

The second potential source of error in NuRDO simulations is in Excel's ability to randomly choose observations belonging to the specified distribution (means and standard deviations). By their very nature when random numbers are involved the results are not perfectly reproducible (especially obvious with small sample sizes). Still, if enough simulations are obtained, the means and standard deviations should closely approach those specified. When one million simulations were made from populations specified to have a mean of 100 and SD of 10, means ranged from 99.9781 to 100.0186 and SD ranged from 9.9846 to 10.0277 (n 10). The overall mean was 99.9976 for 0.0024% error (n 10⁷). The overall SD was 9.9997 for 0.0265% error (n 10⁷).

Therefore, the errors inherent in NuRDO's implementation are much smaller than the experimental error in means and standard errors in typical nutrition experiments (Tables 1–3). Experimental error should not be expected to cause noticeable differences in model parameter estimates. The key to using NuRDO efficiently is to have a good estimate of expected variation under the experimental conditions being simulated, and to perform enough simulations that the coefficient estimates have acceptable variation for the intended use of the experimental results.

The NuRDO workbook was designed to help nutritionists to design experiments using several common models that all may accurately describe nutrition response kinetics under particular circumstances⁽¹¹⁾. The comparisons made in Tables 1–3 and Fig. 1–4 are mainly based on the BLL and BLQ models, as they are the ones that feature requirement estimates with standard errors. The standard errors of the requirements are direct indications of the precision that may be expected from an actual experiment. Of course, that is if the experiments that are actually performed have subject to subject variation and RC similar to past observations. NuRDO could be used in a similar fashion to estimate the precision in estimating the R_{max} or K_m for the SK model⁽¹⁾ or the nutrient concentration giving the maximum response in the QP model. We implemented NuRDO with four commonly used nutritional response models assuming that there is no model appropriate for all situations and researchers should further adapt it to other situations as they deem

appropriate. Nutritional kinetics have been studied in a wide variety of species including humans rats, mice chicks and fish^(1–8). The models should be appropriate for all organisms since enzyme catalysed reactions are common processes of all living things. The NuRDO workbook was developed with four common response models, but is easily adaptable to simulate data from practically any other response model considered more appropriate for any organism.

Acknowledgements

This research was funded by Georgia Agricultural Experiment Station and Hatch project TEX0-1-9258.

The authors' contributions are as follows: G. M. P., D. V. and R. A. A. designed the workbook; D. V. programmed the workbook; G. M. P. conducted the simulations and wrote the manuscript with input from all authors; G. M. P. had primary responsibility for the final content. R. W. formatted the manuscript; all authors read and approved the final version of the manuscript.

The authors have no financial or personal conflicts of interest to declare.

References

1. Morgan PH, Mercer LP & Flodin NW (1975) General model for nutritional responses of higher organisms. *Proc Natl Acad Sci* **72**, 4327–4331.
2. Almqvist HJ (1953) Interpretation of amino acid requirement data according to the law of diminishing returns. *Arch Biochem Biophys* **44**, 245–247.
3. Baker DH (1986) Problems and pitfalls in animal experiments designed to establish dietary requirements for essential nutrient. *J Nutr* **116**, 2339–2349.
4. Hua K (2013) Investigating the appropriate mode of expressing lysine requirement of fish through non-linear mixed model analysis and multilevel analysis. *Br J Nutr* **109**, 1013–1021.
5. Rakhshandeh A, Htoo JK, Karrow N, *et al.* (2014) Impact of immune system stimulation on the ileal nutrient digestibility and utilisation of methionine plus cysteine intake for whole-body protein deposition in growing pigs. *Br J Nutr* **111**, 101–110.
6. Remø SC, Hevrøy EM & Olsvik PA (2014) Dietary histidine requirement to reduce the risk and severity of cataracts is higher than the requirement for growth in Atlantic salmon smolts, independently of the dietary lipid source. *Br J Nutr* **111**, 1759–1772.
7. Andersen SM, Holen E, Aksnes A, *et al.* (2013) Dietary arginine affects energy metabolism through polyamine turnover in juvenile Atlantic salmon (*Salmo salar*). *Br J Nutr* **110**, 1968–1977.
8. Stephens TV, Payne M, Ball RO, *et al.* (2015) Protein requirements of healthy pregnant women during early and late gestation are higher than current recommendations. *J Nutr* **145**, 73–78.
9. Vedenov D & Pesti GM (2008) A comparison of methods of fitting several models to nutritional response data. *J Anim Sci* **86**, 500–507.
10. SAS Institute Inc. (2004) *SAS User's Guide*. Version 9.1 ed. Cary, NC: SAS Institute Inc.
11. Pesti GM, Vedenov D, Cason JA, *et al.* (2009) A comparison of methods to estimate nutritional requirements from experimental data. *Br Poult Sci* **50**, 16–32.