

# Method: Using generalized additive models in the livestock animal sciences

G. L. Simpson<sup>a</sup>

<sup>a</sup>Aarhus University, Department of Animal and Veterinary Sciences, Blichers Allé 20, Tjele, Denmark, 8830

---

## Abstract

Nonlinear relationships between covariates and a response variable of interest are frequently encountered in animal science research. Within statistical models, these nonlinear effects have, traditionally, been handled using a range of approaches, including transformation of the response, parametric nonlinear models based on theory or phenomenological grounds (e.g., lactation curves), or through fixed degree spline or polynomial terms. If it is desirable to learn the shape of the relationship of one or more covariates and a response, then generalized additive models (GAMs) are an excellent alternative to these traditional approaches. GAMs extend the generalized linear model such that the linear predictor includes one or more smooth functions, parameterised using penalised splines. A wiggleness penalty on each function is used to avoid over fitting while estimating the parameters of the spline basis functions to maximise fit to the data. Modern GAMs include automatic smoothness selection methods to find an optimal balance between fit and complexity of the estimated functions. Because GAMs learn the shapes of functions from the data, the user can avoid forcing a particular model to their data. Here, I provide a brief description of GAMs and visually illustrate how they work. I then demonstrate the utility of GAMs on three example data sets of increasing complexity, to show i) how learning from data can produce a better fit to data than that of parametric models, ii) how hierarchical GAMs can be used to estimate growth data from multiple animals in a single model, and iii) how hierarchical GAMs can be used for for-

mal statistical inference in a designed experiment of the effects of exposure to maternal hormones on subsequent growth in Japanese quail. In each case I show how the fitted model can be used to estimate biologically-motivated values (e.g., persistence of the lactation curve). The examples are supported by R code that demonstrates how to fit each of the models considered, and reproduces the results of the statistical analyses reported here. Ultimately, I show that GAMs are a modern, flexible, and highly usable statistical model that is amenable to many research problems in animal science.

5 **Keywords:** Generalized additive model, Penalised spline, Hierarchical model,  
6 Conditional effects, Basis function

---

## 7 **Implications**

8 Nonlinear relationships between covariates and a response variable of interest are fre-  
9 quently encountered in animal science research. Generalized additive models with au-  
10 tomatic smoothness selection via penalised splines provide an attractive, flexible, data-  
11 driven statistical model that is capable of estimating these relationships. I provide a  
12 description of the generalized additive model and demonstrate its use on three typical  
13 data examples; i) a lactation curve from a dairy cow, ii) growth curves in commercial pigs,  
14 and iii) an experiment on the effects of maternal hormones on growth rates in Japanese  
15 quail.

---

\*Corresponding author  
Email address: [gavin@anivet.au.dk](mailto:gavin@anivet.au.dk) (G. L. Simpson)  
Preprint submitted to *Animal – Open Space*

16 **Specification table**

Subject	Livestock farming systems
Specific subject area	Quantitative analysis of animal performance, growth, and modelling
Type of data	Table, graph, code
How data were acquired	Lactation curve: unstated in original source. Pig growth data: weight measurements derived from a depth camera (iDOL65, dol-sensors a/s, Aarhus, Denmark) and a YOLO (you only look once) algorithm. Quail growth experiment: body mass recorded using a digital balance.
Data format	Lactation curve: processed (averaged). Pig growth data: processed (averages of multiple depth camera-based weight measurements). Quail growth experiment: Raw.
Parameters for data collection	Lactation curve: daily fat content of milk for a single animal (cow 7450). Pig growth data: Pigs raised under conventional husbandry conditions at two farms. Quail growth experiment: 158 eggs from adult Japanese quails provided by Finnish private local breeders were injected with one or a combination of maternal hormones or a saline solution control, and those eggs that hatched successfully were monitored for 78 days for a range of parameters including body mass.
Description of data collection	Lactation curve: fat content of daily milk production was measured for a single cow. Pig growth data: a depth camera observed the pigs <i>in situ</i> and a computer algorithm converted the digital imagery of individual animals into weight estimates. Quail growth experiment: the body mass of hatched quail was recorded at 12 hours post hatching, once every three days for days three to 15, and weekly thereafter until day 78.
Data source location	Lactation curve: unknown. Pig growth data: Data from two farms were reported in the original study: A commercial farm in Gronau, Germany, and the experimental farm of the Department of Animal and Veterinary Sciences, Aarhus University, Viborg, Denmark. Quail growth experiment: Finland
Data accessibility	Repository name: Zenodo Data identification number: 10.5281/zenodo.15777270 Direct URL to data: <a href="https://doi.org/10.5281/zenodo.15777270">https://doi.org/10.5281/zenodo.15777270</a>

17

## 18 Introduction

19 Many research questions in the animal sciences involve the estimation of nonlinear  
20 relationships between covariates and a response variable of interest. A classic exam-  
21 ple, with a long history of statistically-based and mathematically-based research, is the  
22 plethora of models that have been described for the estimation of lactation curves from  
23 test day data or automatic milking machines (e.g., [Macciotta et al., 2011](#)). Another, is the  
24 estimation of growth curves in the context of breeding and genetics (e.g., [White et al.,](#)  
25 [1999](#)).

26 Generalized additive models (GAMs, [Hastie and Tibshirani, 1990](#); [Wood, 2017](#)) are a  
27 powerful and flexible class of regression models, which extend the generalized linear  
28 model (GLM) to allow the effects of covariates on the response to be modelled using  
29 smooth, nonlinear functions. Despite the frequency with which such nonlinear relation-  
30 ships are encountered, surprisingly little use of GAMs has been seen in animal science  
31 to date. Notable exceptions include [Hirst et al. \(2002\)](#), [Yano et al. \(2014\)](#), [Benni et al.](#)  
32 [\(2020\)](#), and [Huang et al. \(2023\)](#). The particular advantage of GAMs over parametric  
33 models is that because GAMs do not force a particular model on to the data that can  
34 adapt to the data themselves, providing a substantially improved estimation of covariate  
35 effects and thence the biological interpretations derived from these.

36 In part, the lack of uptake of GAMs in animal science reflects a traditional statistics  
37 workflow grounded in linear mixed effects modelling. Statistical training rarely includes  
38 more advanced models like GAMs, and GAMs are sufficiently different an approach  
39 that many researchers may be wary of using them because they are unfamiliar with the  
40 nomenclature used to describe the models or the software used to fit them. Where GAMs  
41 have been used in animal science, best practice is often not followed, for example in the  
42 choice of smoothness selection method, or failing to correctly specify the conditional  
43 distribution of the response (e.g. [van Lingen et al., 2023](#)).

44 More attention has been paid to the use of spline terms within linear models, frequently  
45 in comparisons against some form of polynomial-based model, especially Legendre poly-  
46 nomials (e.g., [Nagel-Alne et al., 2014](#); [Silvestre et al., 2006](#); [Macciotta et al., 2010](#); [Brito  
47 et al., 2017](#)). There, the focus has largely been on choosing the number of knots in  
48 the spline basis expansion and on the placement of those knots. Modern GAMs largely  
49 make such choices redundant; with penalised splines, a wiggleness penalty is used to  
50 avoid over-fitting, and low-rank eigen bases, such as the low-rank thin plate regression  
51 spline basis of [Wood \(2003\)](#), avoid the knot-placement issue for most problems.

52 A guide to the use of GAMs in the animal science setting, written with users in mind,  
53 is needed to raise awareness of the utility of these flexible models and to promote best  
54 practice in their fitting. Below, I describe GAMs and demonstrate visually how they work.  
55 Then I apply GAMs to three different examples that scale in complexity and which repre-  
56 sent typical data problems encountered in animal science. As well as showing the pre-  
57 dictive ability of GAMs, the examples demonstrate how the fitted models can be used  
58 to provide estimates that are biologically relevant, such as treatment differences, esti-  
59 mates of growth rates, or persistency in milk production. The examples are supported  
60 by tutorials that contain the computer code needed to fit the models in the R statistical  
61 software ([R Core Team, 2025](#)), which are available alongside the source code for the  
62 manuscript itself.

## 63 **Materials and methods**

64 I begin with an introduction to GAMs and how they work. This section is necessarily  
65 general to allow for a brief summary of the main points and illustration of how penalised  
66 splines work. In the subsequent section describing the three worked examples I provide  
67 more case-specific descriptions of the models that are fitted to data.

### 68 *Generalized additive models*

69 A basic GAM ([Hastie and Tibshirani, 1990](#)) has the following form

$$\begin{aligned}
70 \quad & y_i \sim \mathcal{D}(\mu_i, \phi) \\
71 \quad & \mathbb{E}(y_i) = \mu_i \\
72 \quad & g(\mu_i) = \mathbf{X}_i \boldsymbol{\gamma} + \sum_j f_j(\bullet), \quad i = 1, \dots, n; \quad j = 1, \dots, J
\end{aligned}$$

73 where  $y_i$  is the  $i$ th observation of a univariate response variable of interest that is mod-  
74 elled as a function of covariates on the scale of a link function  $g(\cdot)$ .  $\mathcal{D}(\mu_i, \phi)$  is a distri-  
75 bution, typically from, though not limited to, the exponential family of distributions, with  
76 mean  $\mu_i$  and scale parameter  $\phi$ .  $\mathbf{X}_i$  is the  $i$ th row of the model matrix of any parametric  
77 terms (including the model intercept or constant term), and  $\boldsymbol{\gamma}$  the associated regression  
78 parameters. The  $f_j$  are  $J$  smooth functions of one or more covariates (independent vari-  
79 ables, e.g. days in milk); I use  $\bullet$  as a placeholder for covariate(s), but in the simplest  
80 case of a smooth of a single continuous covariate we have  $f_j(x_{ij})$ , a univariate smooth  
81 where  $x_{ij}$  is the  $i$ th observation of the  $j$ th covariate. For a smooth of two or more contin-  
82 uous variables, a tensor product smooth representing smooth main effects and smooth  
83 interaction, the linear predictor would include  $f(x_{i1}, x_{i2})$ , for example.

84 In the remainder of this section, I aim to present *just enough* detail about GAMs to af-  
85 ford the reader a general understanding of what is involved in estimating such a model  
86 so that they can appreciate how GAMs work, and how we aim to avoid overfitting or  
87 having to choose how complex each smooth function should be. The original approach  
88 (Hastie and Tibshirani, 1990) for fitting GAMs, known as *backfitting*, required the ana-  
89 lyst to specify how many degrees of freedom each function in the model should take *a*  
90 *priori*; this was viewed by many as being too subjective for more than exploratory anal-  
91 ysis. With modern automatic smoothness selection methods this problem has largely  
92 been resolved through the use of penalised splines and fast algorithms for smoothness  
93 selection (e.g., Wood, 2011; Wood et al., 2016).

#### 94 *Penalised splines*

95 In a GAM, the smooth functions  $f_j()$  are typically represented in the model using splines,  
96 although other functions fit into this framework, most notably iid Gaussian random effects.  
97 A spline is composed of  $K$  basis functions,  $b_k()$ , and their associated coefficients,  $\beta_k$

$$f_j(x_{ij}) = \sum_{k=1}^K \beta_k b_k(x_{ij}) \quad (1)$$

98 where the basis functions  $b_k()$  could be piecewise cubic polynomials for a cubic regres-  
99 sion spline, but more generally are just the component functions of a spline (e.g. B  
100 splines, thin plate splines). For identifiability reasons, the basis is subject to a sum-to-  
101 zero constraint to remove the constant function from the span of the basis, which allows  
102 inclusion of a separate constant term or intercept in the model; this is desirable if we also  
103 want to include categorical (factor) terms in the model for example.

104 Figure 1a shows a cubic B spline ( $K = 10$  basis functions) fitted to simulated data. Fit-  
105 ting a spline involves finding estimates for the coefficients of basis functions,  $\beta_k$ . These  
106 coefficients weight the individual basis functions as shown in Figure 1a. To find the value  
107 of the fitted spline at each value of  $x$ , we sum up the values of the weighted basis func-  
108 tions evaluated at each value of  $x$ . This yields the light blue curve in Figure 1a. The  
109 coefficients for the basis functions,  $\beta_k$ , are determined by forcing the fitted function to  
110 go as close to the data as possible. As shown in Figure 1a, we largely recover the true  
111 function from which the data were simulated. The B spline basis functions for a covariate  
112  $x$  after the absorbing the sum-to-zero constraint into the basis are shown in Figure 1b.

113 Expanding  $x_j$  into many basis functions in this way means we must be cognizant of  
114 the risk of over fitting our sample of data. Taken to the extreme, we could obtain an  
115 arbitrarily close fit to the data by using as many basis functions as there are data (i.e.,  
116  $K = n$ ), but all this would achieve in practice is the replacement of the data with a set of  
117 coefficients.

118 This raises the question of how many basis functions should be used? If there were  $n$   
 119 unique (i.e., no duplicated values) values of the covariate  $x$  in the data set, one option  
 120 would be to use  $n$  basis functions, leading to a situation where our model would have as  
 121 many coefficients as unique data. We gain very little however by using  $n$  basis functions  
 122 from a statistical viewpoint. In practice, we typically use  $\mathcal{N} \ll n$  basis functions. Yet,  
 123 even with just  $\mathcal{N}$  basis functions, we face the real possibility that our model may overfit  
 124 the data if  $\mathcal{N}$  is too large. To address this, modern approaches to fitting GAMs use pe-  
 125 nalisised splines with automatic smoothness selection (e.g. [Eilers and Marx, 1996](#); [Wood,](#)  
 126 [2011](#); [Wood et al., 2016](#)).

127 If our aim is to avoid overfitting, we need to penalise highly complex fitted functions,  
 128 and hence need to define what we mean by “complex”. A complex function would be  
 129 one that “wiggles” about markedly as we move from low to high values of the covariate,  
 130  $x$ . Such wiggling around implies that the function has a high amount of curvature, which  
 131 we measure using the second derivative of the fitted function. While there are other  
 132 definitions for complexity that we can use, this is the typical measure used for penalized  
 133 splines in statistical models. Therefore, to measure the wiggleness of a fitted function we  
 134 need to integrate (sum up) the second derivative of the function  $f$  over the range of  $x$

$$\int_x f''(x)^2 dx = \beta^T \mathbf{S} \beta \quad (2)$$

135 where  $f''$  indicates the second derivative of  $f$ , and note that we are integrating the *square*  
 136 of the second derivative because we need to allow for both negative and positive curva-  
 137 ture as the function wiggles. Conveniently, we can compute this integral as a function  
 138 of the model coefficients  $\beta$  as shown on the right hand side of Equation 2.  $\mathbf{S}$  is a known  
 139 penalty matrix, which encodes the complexity of the basis functions. The penalty matrix  
 140 for the basis expansion shown in Figure 1a is displayed in Figure 1c. Typically, the soft-  
 141 ware used to fit the GAM, such as the R package *mgcv* that is used here, will calculate  
 142  $\mathbf{S}$  for each spline in the model.



Figure 2 shows three different estimated smooths for the simulated data shown in Figure 1, plus their wiggleness or integrated squared second derivative. Fitting a GAM requires us to balance the fit to the data *and* the complexity of the resulting model; we wish to avoid both over fitting the data (Figure 2a) and over smoothing (Figure 2c) the data. We want the fitted functions to be just “wiggly enough” to approximate the true, but unknown functions (Figure 2b).

To that end, we find estimates of the basis function coefficients,  $\beta_k$ , that make the fitted spline go as close to the data as possible, without overfitting; in practice, this is done by maximising the penalised log likelihood

$$\ell_p(\beta, \lambda) = \ell(\beta) - \frac{1}{2\phi} \sum_j \lambda_j \beta_j^T \mathbf{S}_j \beta_j ,$$

where  $\ell_p(\beta)$  is the penalised log likelihood and  $\ell(\beta)$  the log likelihood of the data, given  $\beta$ , and the remainder is the wiggleness penalty for the model. The log likelihood ( $\ell(\beta)$ ) measures how well our model fits the data, while the wiggleness penalty  $\beta_j^T \mathbf{S}_j \beta_j$  measures how complex the model is. Note that any parametric coefficients,  $\gamma$ , including the intercept have been absorbed into  $\beta$  for convenience. The  $\lambda_j$  are known as the smoothing parameters of the model, and it is these smoothing parameters that actually control how much the wiggleness penalty affects the  $\ell_p(\beta)$ . We can think of the  $\lambda_j$  as tuning or hyper-parameters of the model.

As mentioned in the introduction, the knot placement problem can largely be averted through the use of a low rank thin plate regression spline basis (Figure 1 panels d–f). The spline basis described previously was a cubic B spline basis, and for that basis the knots were placed at evenly-spaced *quantiles* of the covariate. In general, the exact positioning of the knots is unimportant, so long as they are spread out over the range of the covariate. If we wish to avoid the knot placement altogether, we could replace the B spline basis with a low rank thin plate regression spline basis (TPRS, Wood, 2003).

168 We start with a basis function at each unique value of the covariate and then transform  
 169 and truncate the entire basis with an eigen decomposition and retain the  $K$  eigenvectors  
 170 with the smallest eigenvalues as  $K$  new basis functions. The decomposition removes  
 171 much of the excessive wiggleness that  $n$  basis functions would provide, while retaining  
 172 many of the good properties of the original basis (Wood, 2003). The main downside of  
 173 the TPRS basis is that it is computationally expensive to form when setting up the model;  
 174 for large data problems, with many thousands of data, simpler bases such as the cubic  
 175 regression spline or B spline may be preferred. In the *mgcv* software used here, the  
 176 TPRS basis is the default basis used for univariate smooths (Wood, 2025, 2011).

177 The algorithms used to fit GAMs need to estimate the model coefficients,  $\beta$ , and choose  
 178 appropriate values of the smoothing parameters,  $\lambda_j$ . This process is known as smooth-  
 179 ness selection, and there are several approaches to smoothness selection that can be  
 180 taken. One is to treat the problem as one of prediction, and choose  $\lambda_j$  in such a way  
 181 as to minimise the cross-validated prediction error of the model. In practice it would be  
 182 computationally costly to actually cross-validate the model for fitting, so we approximate  
 183 the prediction error by minimising the generalised cross validation (GCV) error, AIC, or  
 184 similar measure. An alternative means of smoothness selection is to take a Bayesian  
 185 view of the smoothing process (see Miller, 2025, for an accessible introduction to this  
 186 viewpoint); in doing so, the  $\beta_j^T S_j \beta_j$  can be viewed as multivariate normal priors on the  
 187  $\beta$ . From this Bayesian view of smoothing, we find that the criterion we wish to minimise  
 188 is that of a mixed effects model. Hence, we can think of the wiggly parts of the  $f_j$  as  
 189 being fancy random effects, while the smooth parts of the  $f_j$  are fixed effects, and the  $\lambda_j$   
 190 are inversely proportional to the random effect variances one would observe if the model  
 191 were fitted as a mixed effects model. The Bayesian approach to smoothing can involve  
 192 fully Bayesian estimation using Markov chain Monte Carlo (MCMC) or simulation free  
 193 estimation via the integrated nested Laplace approximation (INLA), or, using the equiva-  
 194 lence of splines and random effects, we can take an empirical Bayesian approach, which  
 195 yields the posterior modes of the  $\beta_j$ , or the maximum a posteriori (MAP) estimates.

### Why not polynomials?

At this stage, it is not unreasonable to question whether we need anything so complicated as a penalized spline; would polynomial terms not work just as well? A polynomial model includes powers of  $x$  in the linear predictor, e.g.,  $\eta_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3 + \dots + \beta_p x_i^p$ . This is also a basis expansion of  $x$ , where the basis functions are the zeroth, first, second, third etc. powers of  $x$ ;  $b_1(x_i) = x_i^0 = 1$ ,  $b_2(x_i) = x_i^1 = x_i$ ,  $b_3(x_i) = x_i^2$ ,  $b_4(x_i) = x_i^3$ , etc. From the point of view of estimating an unknown function, using this polynomial basis expansion poses two significant problems.

First, while Taylor's theorem suggests that polynomials are useful in approximating an unknown function *at a specific point* (e.g., value of  $x_i$ ), this doesn't imply that polynomials will be similarly useful when interest is on the entire domain of  $x$ . If we consider the simpler problem of interpolating data, polynomials tend to perform poorly when approximating functions while requiring the curve to pass through the data points and maintain continuous derivatives with respect to  $x$ ; to meet these requirements the fitted polynomial curve will tend to oscillate wildly, especially toward the edges of the covariate space (e.g. [Wood, 2017](#), §4.2, pp. 162–4). This problem persists in the regression model setting where we are interested in smoothing data contaminated with noise.

Second, while estimating the polynomial model is trivial, model selection becomes an issue. Model selection is needed to determine the order of polynomial term used in the model as this will ultimately determine the wiggleness of the fitted function. This could be achieved by use of backward selection, while penalized regression with ridge or lasso penalties would shrink specific coefficients thus reducing model complexity. Performing statistical inference (forming confidence intervals, computing  $p$  values, etc.) after conducting these model selection procedures is difficult and an active area of research in statistics. Ultimately, this means such approaches tend to be used where the predictive ability of the model is the primary focus. The theory of GAMs is sufficiently advanced that statistical inference is possible currently.

223 A further problem with polynomial bases is that they are non-local; the fitted value of  
224  $y$  at some value  $x = x_0$  is very sensitive to the values of  $x$  a long way away from  $x_0$ ,  
225 especially where the order of polynomial  $p$  is greater than 2 or 3.

## 226 *Examples*

227 In the remainder of the section I describe three representative examples that demon-  
228 strate the utility of GAMs to address problems in animal science.

### 229 *Lactation curves*

230 As a simple illustration of the benefits of GAMs to learn the functional form of a rela-  
231 tionship between a response variable and a covariate, rather than impose one through  
232 a parametric model, I reanalyse a small data set of average daily fat content per week of  
233 milk from a single cow (Figure 3). The data were reported in [Henderson and McCulloch](#)  
234 [\(1990\)](#). Initially, I followed their [\(Henderson and McCulloch, 1990\)](#) analysis and fitted  
235 a gamma generalized linear model (GLM) with log link function. However, subsequent  
236 analysis of this model (and the GAM alternative described below) showed that the aver-  
237 age daily fat data are under dispersed relative to that assumed by the gamma distribution.  
238 Instead, a Tweedie GLM (log link) was fitted, which has the same response shape as  
239 the gamma GLM. The Tweedie distribution is a family of distributions with support on  
240 the non-negative real numbers ([Smyth, 1996](#)), making it especially suitable for positive  
241 or non-negative continuous random variables, such as milk yield, or milk fat content as  
242 observed in this example. The Tweedie distribution contains the gamma distribution as  
243 a special case, but has more flexibility in terms of the mean-variance relationship as-  
244 sumed of the response; the gamma has a quadratic mean-variance relationship, while  
245 those admitted by the Tweedie distributions considered here cover the range from linear  
246 (i.e. Poisson-like) to quadratic (i.e., gamma-like) mean-variance scaling.

247 The Tweedie GLM fitted had the following form

248

$$\mathbf{fat}_i \sim \mathcal{T}(\mu_i, \phi)$$

249

$$\log(\mu_i) = \beta_0 + \beta_1 \log(\mathbf{week}_i) + \beta_2 \mathbf{week}_i .$$

250

[Henderson and McCulloch \(1990\)](#) compared the fit of the GLM model with several other formulations and models, including the model of [Wood \(1967\)](#)

251

252

$$\mathbf{fat}_i = \alpha \mathbf{week}_i^{\delta} \exp(\kappa \mathbf{week}_i) + \varepsilon_i$$

253

where  $\alpha$ ,  $\delta$ , and  $\kappa$  are parameters whose values are to be estimated, and  $\varepsilon$  is a Gaussian error term. Note that  $\kappa$  is usually presented as  $-\kappa$ , hence the parameter estimate takes values that are the negative of the usual values. The linear predictors in the GLM and Wood's model are equivalent because of the log link function;  $\beta_0 = \log \alpha$ ,  $\beta_1 = \delta$ , and  $\beta_2 = \kappa$ . However, we would not expect both models to produce the same fitted lactation curve as different distributional assumptions are being made; in the GLM version we assume the average daily fat values are conditionally Tweedie distributed, while in Wood's model they are assumed to be conditionally Gaussian distributed.

261

These models were compared with a GAM version of the log-link, Tweedie GLM, where the fixed functional form of the lactation curve has been replaced by a smooth function of  $\mathbf{week}_i$ ,  $f(\mathbf{week}_i)$ , that is to be estimated from the data

262

263

264

$$\mathbf{fat}_i \sim \mathcal{T}(\mu_i, \phi)$$

265

$$\log(\mu_i) = \beta_0 + f(\mathbf{week}_i)$$

Wood's (1967) model was estimated via nonlinear least squares using the `nls()` function in R (version 4.5.0, R Core Team, 2025). The Tweedie GLM and GAM were fitted using the `gam()` function of the *mgcv* package (version 1.9.3, Wood, 2025, 2011) for R.  $K = 9$  basis functions were used for  $f(\text{week}_i)$  after application of the identifiability constraint.

The fitted models were subsequently used to estimate some biologically-relevant parameters that could be used to inform breeding or optimise production methods. The week of peak milk fat content, the estimated fat content at the peak, and the rate of decline in the lactation curve at a point midway between the peak and the end of lactation are estimated for all models. For Wood's model, the timing of peak fat content is given by  $\delta / -\kappa$  and peak fat content by  $\alpha(\delta / -\kappa)^\delta e^{-\delta}$  (Wood, 1967).

For the Tweedie GAM, we do not have such directly interpretable parameters, therefore the peak is identified by evaluating the model at a grid of 1000 values over the  $\text{week}_i$  covariate and noting the point at which the estimated fat content is maximised. This process also yields the estimated maximum fat content. While the Tweedie GLM has the same interpretable parameters as Wood's model, determining the uncertainty in the estimate time to peak fat content and peak fat content is not trivial. Instead, the same procedure as that outlined for the GAM was also followed for the GLM.

The first derivative (slope) of the falling limb of a lactation curve is often used as a measure of persistency in dairy animals. While there are several such derivatives that can be derived directly from the parameters of Wood's model (e.g., Dijkstra et al., 2010; Oliveira et al., 2020), exact equivalents for the GAM may not be simple to implement. Instead, for illustration and to avoid differences between models due to the specific details of implementation, for all three models I estimate the first derivative of the lactation curve at a point midway between the peak and the end of lactation. This is intentionally similar to the relative decline measure of Dijkstra et al. (2010). For all three models the first derivative is calculated using a central finite difference with  $h$ , the step size, set to be

1 day. The central finite difference estimate of the first derivative of the lactation curve,  $\hat{f}'$ , is

$$\hat{f}' = \lim_{h \rightarrow 0} \frac{f\left(x + \frac{h}{2}\right) - f\left(x - \frac{h}{2}\right)}{h}.$$

The uncertainty in each of the three quantities estimated from the models was computed using 10,000 parametric bootstrap samples in the case of Wood’s model using the `nlsBoot()` function from package *nlstools* (version 2.1.0, [Baty et al., 2015](#)), while for the GLM and GAM 10,000 draws from the posterior distribution of the parameters  $\beta$  using a Gaussian approximation to the posterior were used (See [Miller, 2025](#), for a brief introduction to this idea). Posterior sampling was performed using the `fitted_samples()` function from package *gratia* ([Simpson, 2024](#)).

### Pig growth

In this second example, I illustrate how individual growth curves can be estimated using various forms of hierarchical GAM. A hierarchical GAM (HGAM, [Pedersen et al., 2019](#)) is a GAM that includes smooths at different hierarchical levels of the data structure; HGAMs are the GAM equivalent of hierarchical or mixed effects models, which are frequently used with such data. In this example, different parameterisations of the model lead to either direct estimation of each individual animal’s growth curve, or a decomposition into an average growth curve plus animal-specific deviations from this average curve. Although previously covered by [Pedersen et al. \(2019\)](#) in detail, a new addition here is the use of a constrained factor smooth to produce animal-specific deviations that are truly orthogonal to the average curve. This basis type was not available to [Pedersen et al. \(2019\)](#), and helps avoid switching to a first derivative penalty for the animal-specific curves that would be required to make the model identifiable if a “factor by” smooth was used.

Automated estimation of body weights is a useful on-farm method for continuously monitoring growth of commercial pigs. I analyse a subset of the weight data reported by

318 [Franchi et al. \(2023\)](#) from a study by [Bus et al. \(2025\)](#), using data from a single pen of 18  
 319 pigs (Figure 5). The body weight data were obtained using a depth camera (iDOL65, dol-  
 320 sensors A/S, Aarhus, Denmark), and the each weight observation is the daily average of  
 321 multiple measurements made by the camera. As the number of weight measurements  
 322 made each day varied per animal and per day, each weight observation in the data is the  
 323 average of a variable number of measurements. To accomodata this, the model included  
 324 the number of measurements averages as an observation weight, where the precision  
 325 of each response data is proportional to the number of measurements averaged.

326 The data exhibit both common (shared) and animal-specific variation. To model these  
 327 features, four different HGAMs were fitted to the pig weight data. All of the models  
 328 ultimately provide animal-specific estimates of weight over time, but they decompose  
 329 the growth curves in different ways. The weight data were assumed to be conditionally  
 330 gamma distributed, with log link,  $\text{weight}_i \sim \mathcal{G}(\mu_i, \phi)$ , with  $\log(\mu_i) = \eta_i$ , where  $\eta$  is the  
 331 linear predictor. The gamma distribution was chosen as it has support on the positive real  
 332 numbers; the weight of a pig is a necessarily positive continuous random variable. Other  
 333 choices for the conditional distribution would be a log-normal or a Tweedie distribution,  
 334 but standard model diagnostics did not indicate any problems with the use of the gamma  
 335 distribution.

336 The linear predictors for each of the four models were:

$$337 \quad \text{P1: } \eta_i = \beta_{a(i)} + f_{a(i)}(\text{day}_i)$$

$$338 \quad \text{P2: } \eta_i = \beta_0 + f_1(\text{day}_i) + f_{a(i)}(\text{day}_i)$$

$$339 \quad \text{P3: } \eta_i = \beta_0 + f_{a(i)}^*(\text{day}_i)$$

$$340 \quad \text{P4: } \eta_i = \beta_0 + f_1(\text{day}_i) + f_{a(i)}^*(\text{day}_i)$$



341 where  $a(i)$  is an indicator function that gives the animal to which the  $i$ th observation  
 342 belongs. Model P1 includes the mean weight of each animal through a parametric fac-  
 343 tor term,  $\beta_{a(i)}$ , plus a smooth of observation day  $\text{day}_i$  *per* animal. The parametric factor  
 344 term is required because the animal-specific smooths in this model are each subject to  
 345 the sum-to-zero-constraint and as such do not contain the constant functions that are  
 346 needed to model the average weight of each animal. These smooths are known infor-  
 347 mally as “factor by smooths”. Model P2 decomposes the data into an average growth  
 348 curve,  $f_1(\text{day}_i)$ , plus smooths, one per animal, that represent deviations from the aver-  
 349 age smooth,  $f_{a(i)}(\text{day}_i)$ . Unlike the “factor by smooths”, these deviation smooths include  
 350 constant terms for each animal's average weight, hence only a constant term,  $\beta_0$  is in-  
 351 clude in the parametric part of this model. In both models P1 and P2, the  $f_{a(i)}(\text{day}_i)$   
 352 smooths each have their own smoothing parameters ( $\lambda$ s), allowing the wiggleness of  
 353 each animal's growth curve to vary, if supported by the data.

354 Model P3 is similar to model P1, except that each animal's growth curve,  $f_{a(i)}^*(\text{day}_i)$ ,  
 355 shares a single smoothing parameter ( $\lambda$ ) for the wiggleness and hence assumes that  
 356 the wiggleness of each curve is similar. These smooths are denoted with a superscript  
 357 \* to indicate the shared smoothing parameter, and can be thought of as the smooth  
 358 equivalent of random slopes and intercepts; informally, we refer to these as “random  
 359 smooths”. These smooths are fully penalised and contain constant terms to model the  
 360 average weight of each animal, hence only the intercept,  $\beta_0$  is included in the parametric  
 361 part of the model. Model P4 is similar to model P2 in terms of the decomposition into an  
 362 *average* smooth and animal-specific smooth deviations, but, like model P3, the animal-  
 363 specific smooths share a smoothing parameter and therefore assumes they have similar  
 364 wiggleness.

365 The range of GAMs fitted to the pig growth data is intended to illustrate the modelling  
 366 choices open to the analyst; how are the individual growth curves modelled, and how  
 367 much variation in wiggleness is allowed per growth curve. The choices then are; is each

368 animal's growth curve modelled directly (P1 and P3) or are the growth curves modelled  
369 as an "average" growth curve plus animal-specific deviations from this (models P2 and  
370 P4), and should the animal-specific curves be allowed to have different wiggleness if  
371 needed (models P1 and P2), or should they all have the same wiggleness (models P3  
372 and P4).

373 For further details on the different approaches used in the models described above, see  
374 [Pedersen et al. \(2019\)](#). Each of the smooths in the models used  $K = 9$  basis functions,  
375 after application of identifiability constraints.

376 The estimated daily growth rates on November 15<sup>th</sup>, 2021 for each pig were calcu-  
377 lated from the selected GAM using a central finite difference estimate of the first deriva-  
378 tive of the fitted growth curve. The estimated growth rate is given by the median of  
379 the posterior distribution and an 95% Bayesian credible interval on the estimate is pro-  
380 vided. This date was chosen as weight observations for three of the pigs were missing  
381 after November 1<sup>st</sup>. All estimates were based on 10,000 draws from the posterior dis-  
382 tribution, using a Gaussian approximation to the posterior implemented in the functions  
383 `response_derivatives()` and `derivative_samples()` from package *gratia*.

#### 384 *Japanese quail*

385 In the third example, I demonstrate how to fit GAMs in the context of a designed exper-  
386 iment to obtain estimated treatment effects.

387 The data ([Sarraude et al., 2020a](#)) are from an experiment into the short- and long-term  
388 effects on quail of elevated exposure to different types of maternal thyroid hormone in  
389 Japanese quail, *Coturnix japonica* ([Sarraude et al., 2020b](#)). Briefly, the yolks of  $n = 57$   
390 eggs were experimentally manipulated by injection with the prohormone,  $T_4$ , its active  
391 metabolite,  $T_3$ , or both  $T_4$  and  $T_3$  ( $T_3T_4$ ), or a saline solution that acted as a control (CO).  
392 Body mass was initially measured 12 hours after hatching. Between days 3–15, body  
393 mass was recorded every three days, then, between days 15–78, once per week, using

394 a digital balance.

395 For the same reasons as in the pig growth example, the quail weight data are nec-  
 396 essarily positive and were provisionally assumed to be conditionally gamma distributed.  
 397 Despite this being a reasonable working assumption, model diagnostics identified devi-  
 398 ations from this assumption, and instead the data were modelled as being conditionally  
 399 Tweedie distributed. The Tweedie family of distributions contains the Poisson (power,  
 400  $p = 1$ ) and gamma distributions ( $p = 2$ ) as special cases, but is more flexible than  
 401 gamma, allowing for a range of mean-variance relationships through the power param-  
 402 eter,  $p$ , of the distribution. In the Tweedie GAMs that were fitted, the power parameter  
 403  $p$  was allowed to vary between 1–2 and was estimated as an additional model constant  
 404 term. Models fitted were:

$$405 \quad \text{Q1: } \eta_i = \beta_0 + f(\text{day}_i) + f_{\text{sex}(i)}(\text{day}_i) + f_{\text{egg}(i)}^*(\text{day}_i) + \xi_{\text{mother}(i)}$$

$$406 \quad \text{Q2: } \eta_i = \beta_0 + f(\text{day}_i) + f_{\text{treat}(i)}(\text{day}_i) + f_{\text{sex}(i)}(\text{day}_i) + f_{\text{egg}(i)}^*(\text{day}_i) + \xi_{\text{mother}(i)}$$

$$407 \quad \text{Q3: } \eta_i = \beta_0 + f(\text{day}_i) + f_{\text{treat}(i)}(\text{day}_i) + f_{\text{sex}(i)}(\text{day}_i) + f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i) \\ 408 \quad + f_{\text{egg}(i)}^*(\text{day}_i) + \xi_{\text{mother}(i)}$$

$$409 \quad \text{Q4: } \eta_i = \beta_0 + f(\text{day}_i) + f_{\text{treat}(i)}(\text{day}_i) + f_{\text{sex}(i)}(\text{day}_i) + f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i) \\ 410 \quad + \psi_{\text{egg}(i)} + \xi_{\text{mother}(i)}$$

411 where  $\text{day}_i$  is the number of days since hatching,  $\text{treat}(i)$ ,  $\text{sex}(i)$ ,  $\text{egg}(i)$ , and  $\text{mother}(i)$   
 412 indicate to which treatment group, sex, bird, and mother the  $i$ th observation belongs.  
 413 Smooth functions with subscript  $\text{treat}(i)$ ,  $\text{sex}(i)$ , or  $\text{egg}(i)$  are factor-smooth interac-  
 414 tions representing deviations from the “average” smooth,  $f(\text{day}_i)$ , for the indicated fac-  
 415 tor.  $f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i)$  represents a higher order factor-smooth interaction, which allows  
 416 the time varying treatment effects to also vary between male or female quail.  $\psi_{\text{egg}(i)}$

417 and  $\xi_{\text{mother}(i)}$  are iid Gaussian random intercepts for individual quail and their mother re-  
 418 spectively. A  $f^*$  represents a random smooth, where each smooth in the set shares a  
 419 smoothing parameter. The factor-smooth interactions in this model were fitted using the  
 420 constrained factor-smooth interaction basis in the *mgcv* package; this basis excludes the  
 421 main effects (and lower-order terms in the case of higher-order interactions) to insure  
 422 that the basis functions are orthogonal to those lower order terms.

423 Model Q1 represents a null model containing no treatment effects but models the re-  
 424 maining features of the data, decomposing the growth curves into an average effect, a  
 425 sex-specific effect, and individual quail-specific effects, plus a maternal effect. Model Q2  
 426 extends Q1 by adding the treatment effect,  $f_{\text{treat}(i)}(\text{day}_i)$ , which models deviations from  
 427 the average curve for each of the four treatment levels. Model Q3 extends Q2 to allow  
 428 different treatment-specific curves for male and female quail. Model Q4 is a variant of  
 429 Q3, replacing the quail-specific growth curve deviations with an individual level random  
 430 intercept. *A priori*, model Q3 represents the complete set of hypotheses an analyst might  
 431 expect to consider for these data.

432 Each smooth in the quail models used  $K = 9$  basis functions after application of iden-  
 433 tifiability constraints, except the quail-specific smooths,  $f_{\text{egg}(i)}^*(\text{day}_i)$ , which used  $K = 6$   
 434 basis functions per individual. This reduced number of basis functions per smooth re-  
 435 flects the fact that after accounting for the average shape of the growth curves, plus sex  
 436 and treatment deviations, any remaining individual-specific variation from these other  
 437 effects would be smaller in magnitude and less complex (wiggly).

438 These data were previously analysed using GAMs by [Sarraude et al. \(2020a\)](#) and [Mo-  
 439 rota et al. \(2021\)](#); [Sarraude et al. \(2020a\)](#) and [Morota et al. \(2021\)](#) use an ARMA(1,1)  
 440 correlation for model residuals to account for the longitudinal nature of the data that re-  
 441 main unmodelled after inclusion of smooths of bird age conditioned on treatment or sex,  
 442 and it is unclear if a non-Gaussian distribution was used. The main advance provided by  
 443 this reanalysis is that instead of treating the individual quail growth curves as stochastic

444 components of the model, here I directly model those curves using “random smooths”.  
445 Additionally, here I assume the response data are conditionally distributed Tweedie, and  
446 I directly assess the treatment by sex interaction.

#### 447 *Smoothness selection & inference*

448 Restricted marginal likelihood (REML) smoothness selection (Wood, 2011) was used to  
449 estimate each of the GAMs fitted to the example data sets. REML was used because it  
450 has slightly better performance in terms of estimating smoothing parameters compared  
451 to marginal likelihood (ML) smoothness selection. We did not use GCV smoothness  
452 selection for these examples because i) prediction error is a less important consideration  
453 here where one is interested in estimation of effects and statistical inference, and ii) GCV  
454 is prone to under smoothing (Reiss and Ogden, 2009; Wood, 2011).

455 In the lactation curve example I use standard model diagnostic techniques (e.g.,  
456 quantile-qauntile (QQ) plots, plots of residuals versus linear predictor values) to identify  
457 the most useful model. In the pig growth example, AIC was used to identify which of  
458 the decompositions of time provided the best fit, but the specific choice of model was  
459 made using domain knowledge.

460 In the quail hormone example, AIC values are reported, but *a priori* I chose to report  
461 results from model Q3, the “full” model from the point of view of the potential hypotheses  
462 under consideration; as the treatment effect was allowed to vary between sexes, I take  
463 an estimation-based approach and quantify the magnitudes of any treatment by sex in-  
464 teractions using the most complex model. The remaining models are fitted and reported  
465 on briefly for illustrative purposes; performing model term selection (for example, test-  
466 ing the higher order  $f_{\text{treat}(i)}(\text{day}_i)$  interaction, and deciding on the basis of a  $p$  value or  
467 AIC whether to retain it in the model or not), would introduce biases into the inference  
468 process. As we currently lack post selection inference methods to deal with this model  
469 form of selection it is prudent to simply not enter into such model selection procedures.

Using model Q3 allows there to be treatment differences between male and female quail; subsequent estimation of values of interest and comparison among these estimates is instead the inference route followed for the quail example. To illustrate, I estimate two quantities of interest: i) the estimated growth rate (slope) of an average quail at  $\text{day} = 20$  in all combinations of treatment and sex, and ii) the estimated weight of an average quail at the end of the experiment ( $\text{day} = 78$ ), again for all combination of treatment and sex. All pairwise comparisons among treatment levels within sex were performed, with adjustment of  $p$  values to control the false discovery rate using the Benjamini–Yekutieli (Benjamini and Yekutieli, 2001) procedure. Estimates and pairwise comparisons were computed using the `slopes()` and `predictions()` functions of the *marginaleffects* package for R (version 0.27.0, Arel-Bundock et al., 2024). Estimates of the expected growth curves for average quail in all combinations of treatment and sex were produced using the `conditional_values()` function in the R package *gratia* (version 0.11.1.9000, Simpson, 2024).

All figures were produced using the R packages *gratia* and *ggplot2* (version 3.5.2, Wickham, 2016).

## Results

### *Lactation curves*

Figure 3a shows the three lactation curves estimated using Wood’s model, a Tweedie GLM, and a Tweedie GAM. While all three models capture the general shape of the lactation data, the Tweedie GLM and, to a lesser extent, Wood’s model, overestimate the peak yield, with the data exhibiting a broader period of peak fat content than is captured by either model. Wood’s model and the Tweedie GLM also fail to capture the features of the mid–late lactation decline in fat content, other than the general decline itself. Conversely, the GAM, as anticipated, estimates a lactation curve that more faithfully tracks the observed data. The model response residuals ( $y_i - \hat{y}_i$ ) for Wood’s model (Figure 3b)

496 and the Tweedie GLM (Figure 3c) show a significant amount of unmodelled signal, while  
497 the response residuals for the Tweedie GAM (Figure 3d) are much smaller and do not  
498 show a residual pattern. Despite using roughly twice as many degrees of freedom as  
499 the other models (Table 1), the Tweedie GAM was clearly favoured in terms of AIC and  
500 root mean squared error of the fitted values.

501 Figure 4 shows the estimated timing of peak fat content, the estimated fat content at  
502 the peak, and the first derivative of the estimated lactation curve at a point midway be-  
503 tween the peak and the end of lactation, for the three estimated models. The peak in fat  
504 content is estimated to be significantly later in the GAM and the timing of the peak much  
505 more uncertain compared to both the fits from Wood's model and the GLM (Figure 4a).  
506 This reflects the observation that Wood's model and the GLM both over-estimate the  
507 fat content in the milk between weeks 5 and 10, and miss the broad period of high fat  
508 content up to week 14. Given the broad peak evident in the data, it is not unsurprising  
509 that the GAM-based estimate of the timing of the peak is much more uncertain than the  
510 other two estimates. The estimated peak fat content in the milk is broadly similar across  
511 the three models, with the GLM fit being slightly higher and more uncertain than either  
512 that for Wood's model or the GAM (Figure 4b). The rate of decay (first derivative) of  
513 the falling limb of the lactation curve is estimated to be substantially faster in both the  
514 GLM and Wood's model compared with the GAM estimate. This discrepancy is due to  
515 clear plateau in fat content between weeks 20 and 27 that is completely missed by the  
516 other two models. It should be noted that the posterior sampling method could equally  
517 be applied to yield an average derivative over the falling limb of the lactation curve if that  
518 was of greater focus.

519 For the timing of the peak and the rate of decay in the lactation curve, the estimates  
520 from the GLM and Wood's model are clearly overly optimistic in their uncertainty; the  
521 results in Figure 4a & c imply quite precisely estimated values and yet these models  
522 exhibit substantial biases in their estimated lactation curves, an issue that largely arises

523 because they force a particular model to the data rather than allowing the data to inform  
524 the shape of the lactation curve.

### 525 *Pig growth*

526 The estimated growth curves for the 18 pigs in the pig growth example are shown in Fig-  
527 ure 5, which were produced from model P2. Of the four models fitted, the two models  
528 that decomposed the growth effects into an average curve plus animal-specific devia-  
529 tions from the average curve, models P2 and P4, resulted in the most parsimonious fits  
530 from the point of view of AIC Table 2. This is due to these two model forms using many  
531 fewer degrees of freedom ( $\sim 64$ ) than either of models P1 (EDF = 83.54) or P3 (EDF =  
532 90.57). Models P1 and P3 do not include the average curve and as a result expend  
533 many more degrees of freedom modelling the same general shape for each animal. In-  
534 terestingly, for these data at least, allowing for a different smoothing parameter (P2) for  
535 each animals' deviation smooth seems to be preferred over using a single smoothing pa-  
536 rameter (P4). In part, this is due to the somewhat idiosyncratic nature of each animal's  
537 growth curve in this data set.

538 Although the deviance explained is very high ( $\sim 96\%$ ) for all models, much of this is  
539 due to use of random effects to model differences between animals, and should not  
540 be taken as a sign that the model can effectively perfectly predict the weight of a pig  
541 under similar conditions; it is clear from Figure 5 that there is much unmodelled variation  
542 around the estimated growth curves. One feature of the data that I do not address here  
543 is the clear variation among animals in the variance of the depth camera-based weight  
544 measurements; animals 5, 6, and 9 in particular, exhibit substantially greater variation  
545 than the other animals in the data set. I return to this point briefly later.

546 Table 3 provides an overview of the two terms in model P2. Although the average curve  
547 ( $f(\text{day}_i)$ ) was allowed to use  $K = 9$  basis functions, the wigginess penalty has shrunk  
548 this back to 7.332 effective degrees of freedom (EDF). The animal-specific deviation  
549 smooths,  $f_{a(i)}(\text{day}_i)$ , were fully penalised and as such used  $K = 10$  basis functions per



550 animal. Here we clearly see the effect of the wiggleness penalty, which has resulted in  
 551 a reduction from a potential 180 EDF for the set of animal-specific smooths to 56.054  
 552 EDF. The test of the null hypothesis for the average growth curve and the omnibus test  
 553 for the pig-specific deviation smooths indicate both are statistically interesting effects.

554 Figure 6a shows the estimated daily growth rate on November 15<sup>th</sup>, 2021 for each of the  
 555 18 pigs. Despite the fact that with GAMs we do not have a simple equation that can be  
 556 used for estimation of biologically-relevant values using the model, posterior simulation  
 557 is a simple and effective way to estimate values of interest, such as the derivatives of  
 558 the response, and provide an uncertainty estimate. For three of the pigs in Figure 6a,  
 559 no weight estimates were available after November 1<sup>st</sup>, so the estimated growth curve  
 560 for these pigs is not constrained by data on these animals, but instead is estimated  
 561 using the shape of the average growth curve ( $f(\text{day}_i)$ ). These estimates do include an  
 562 animal-specific component, however, because we see quite differently-shaped posterior  
 563 distributions for pig 17 compared with pigs 2 and 13, with the estimated growth curve of  
 564 pig 17 being considerably more uncertain than either of the curves for the other two pigs.  
 565 This is reflected in the much broader posterior distribution for pig 17 compared to those  
 566 for pigs 2 and 13 Figure 6b.

#### 567 *Quail hormone experiment*

568 Warning in estimate.theta(theta, family, y, mu, scale = scale1, wt = G\$w, :  
 569 step failure in theta estimation  
 570 Warning in estimate.theta(theta, family, y, mu, scale = scale1, wt = G\$w, :  
 571 step failure in theta estimation

572 As assessed by AIC, models containing treatment effects resulted in no improvement  
 573 in the model fit (Table 4), a result that is consistent with the findings of [Sarraud et al.](#)  
 574 (2020a). However, the results of model Q3 are reported below as this model is consistent  
 575 with the *a priori* assumption that treatment effects might be present, possibly varying by  
 576 sex, which underlay the original experiment. If we were to choose either of models Q1 or

577 Q2 on the basis of AIC, we would be engaging in post-selection inference; having used  
 578 the same data to both fit and test the model, any tests we would wish to conduct on the  
 579 chosen model would yield anti-conservative  $p$  values, and standard errors and credible  
 580 intervals that would be too small or narrow respectively. Furthermore, choosing models  
 581 Q1 or Q2 would in effect be claiming that some or all treatment effects were exactly equal  
 582 to 0 in size; this is implausible and also inconsistent with the data we have observed. It  
 583 is therefore better statistical practice to fit the model containing the effects hypothesised  
 584 *a priori*, and report the estimated effect sizes of those effects.

585 Figure 7 shows the original data and the estimated growth curves for each individual  
 586 quail that were obtained from model Q3. The need for quail-specific random smooths  
 587 is clear enough from the data; model Q4 had the same model structure as that of Q3,  
 588 except for the replacement of the quail-specific random smooths with quail-specific in-  
 589 tercepts, which resulted in an increase in AIC of over 500 units (Table 4).

590 To focus on the estimated treatment effects, model Q3 was evaluated at 100 evenly-  
 591 spaced values over the time covariate (the number of evaluation points was chosen to  
 592 obtain a visually smooth representation of the estimated functions), plus all combinations  
 593 of treatment level and sex. The effects of the quail-specific random smooths and the  
 594 maternal random effects were excluded from these estimates. This results in estimated  
 595 treatment effects for the average quail, which, strictly speaking should not be interpreted  
 596 as population level effects due to the non-identity link function. The resulting estimated  
 597 effects are shown in Figure 8. While there are clear differences in the growth curves  
 598 based on sex, there appears to be little qualitative difference in the effects of treatment  
 599 levels on quail growth.

600 The statistical summary of model Q3 is shown in Table 5. The omnibus tests of the  
 601 treatment specific deviations (row 3;  $f_{\text{treat}(i)}(\text{day}_i)$ ), and the treatment by sex deviations  
 602 from the average curve (row 5;  $f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i)$ ), both have  $p > 0.05$ , which further rein-  
 603 forces the previously made observations that the effects of exposure to maternal thyroid

hormone, if any, are small relative to the uncertainty in the model itself. The  $p$  value for the omnibus test of the quail-specific deviation smooths ( $f_{\text{egg}(i)}(\text{day}_i)$ ) is also greater than 0.05. This is somewhat surprising, given the magnitude of the difference in AIC that is observed when these quail-specific random smooths are replaced by a simple random intercept term ( $\Delta\text{AIC} = 527.388$ ). Despite not achieving “statistical significance”, removing these terms on the basis of  $p$  values would lead to invalid inference, which is to be avoided.

The largest contribution to the overall EDF of this model ( $\text{EDF} = 274.755$ ) is from the quail-specific deviation smooths ( $\text{EDF} = 224.064$ ), although because there are 57 individual animals and the random smooths include random intercepts to account for differences in the average weight of each quail, this is unsurprising. Again, we see the effect of the wigglyness penalty, which has penalised the smooths back to  $\text{EDF} = 224.064$  from a maximum EDF of 342.

To formally assess the differences among treatment effects, pairwise comparisons of treatment levels within sex were conducted for the slope of the average growth curve at day 20. The estimated differences in the slopes of the growth curves are shown in Table 6. For all comparisons, despite some observed differences in the growth rates at day 20 among the treatment levels, the 95% credible intervals include zero for all comparisons. Therefore, if these results reflect the broader population of quail, we should expect that any differences in growth rate due to maternal hormones are small, and largely indistinguishable from zero.

Finally, pairwise comparisons of quail weight at the end of the experiment (day 78) among treatment levels within sex were used to examine differences in estimated weight of quail due to exposure to maternal thyroid hormones. The results of these comparisons are shown in Table 7. The largest observed difference ( $\sim 10.1$  g) in weight of an average quail was between the  $T_4$  hormone treatment and controls in female quail. This represents approximately a 5% decrease in the weight of an average female quail when

631 exposed to the  $T_4$  hormone compared with the untreated controls. Yet, given the uncer-  
632 tainty in the estimates of the model coefficients, the 95% credible interval includes 0 for  
633 this, and all other, comparisons.

#### 634 **Author's Points of View**

635 The results presented above demonstrate the utility of GAMs for modelling the kinds  
636 of data commonly encountered in research involving animals. GAMs are often viewed  
637 unfavourably as being subjective — requiring the user to specify how complex they want  
638 the estimated smooth functions to be — and data hungry, and more of a data visualisation  
639 tool than one suited to formal statistical inference. These views are largely out-dated with  
640 respect to the modern approach to GAMs described here. The subjectivity critique has  
641 been addressed through developments in automated smoothness selection. None of  
642 the data sets analysed here are especially large by today's standards, and where data  
643 sets are much larger, significant research effort has been (e.g., [Li and Wood, 2020](#)) and  
644 continues to be expended to adapt the algorithms to ever higher dimensional problems.  
645 The `bam()` function in the *mgcv* package, for example, can handle data on the order of  
646 millions of rows, with tens of thousands of model parameters, given sufficient availability  
647 of computer memory.

648 As demonstrated in the quail maternal hormone example, formal statistical inference  
649 is entirely feasible. If we ignore the selection of smoothness parameters, the models  
650 described here are little more than generalized linear (mixed) models (GL(M)Ms) once  
651 the basis expansions of covariates have been performed. Modern software tools like the  
652 *marginalEffects* package for R allow a consistent interface to statistical inference across  
653 many disparate statistical models, and GAMs are no different in this regard. Posterior  
654 sampling can also be applied to GAMs to provide point estimates and uncertainties on  
655 any quantity of interest that can be computed using only the model predictions. This  
656 sampling-based approach to inference is incredibly powerful and extremely flexible and

657 complements more standard inference such as the comparisons of contrasts used in the  
658 quail example.

659 The distinct advantage of GAMs over GL(M)Ms is their ability to learn the shapes of  
660 nonlinear relationships between covariates and the response from the data themselves.  
661 This relieves the analyst from having to force data to fit a particular theoretical model,  
662 unless they have a good justification via domain knowledge to use that model. GAMs  
663 also avoid the model selection problems and erratic estimates inherent to modelling with  
664 polynomial basis expansions.

665 The main disadvantage of GAMs is that they are more complex to fit than GL(M)Ms  
666 and the analyst has some additional data modelling choices to make. The main choice  
667 is the number of basis functions,  $K$ , that should be used by each smooth in the model.  
668 The general advice here is for the analyst to imagine the largest amount of wiggleness  
669 that they would expect and then set  $K$  a little larger than this. Of course, the novice  
670 user of GAMs will lack the experience required to do this easily, but for the sorts of  
671 data problems exemplified above, we would not expect highly complex smooth functions,  
672 and the default of  $K = 10$  for univariate smooths in *mgcv* is usually sufficient for many  
673 problems. The key requirement is that the initial number of basis functions used should  
674 be large enough such that the span of functions representable with that basis will contain  
675 the true but unknown function or a close approximation to it. The basis dimension must  
676 be checked of course, and this adds another step to the model appraisal or checking  
677 procedure. With *mgcv* for example, the `k.check()` function provides a test for sufficiency  
678 of the basis size used to fit the model (Pya and Wood, 2016).

679 A further disadvantage is that statistical inference is somewhat more approximate than  
680 with GL(M)Ms. While GAMs share with GL(M)Ms the property of having asymptotically  
681 correct  $p$  values, the  $p$  values for smooths are more approximate than for terms in a  
682 GL(M)M because the current theory on which these tests are based does not account  
683 for the selection of smoothing parameters; for the purposes of the tests, the smoothing

parameters are treated as being fixed and known, but instead they are estimated from the data (Wood, 2013a). While there has been some progress in adapting the theory to include this additional source of uncertainty (e.g., Wood et al., 2016), as yet this has not been used to correct the  $p$  values of tests of smooths. That said, we should avoid making modelling decisions on the basis of  $p$  values, to this additional approximation should not pose a problem in practice.

Finally, GAMs can require more substantial amounts of computing resources to fit than GL(M)Ms once data sets get above tens of thousand observations or where models include several or complex random effect terms, using the algorithms provided by *mgcv*. The main requirement is computer memory, although the `bam()` function can help with this using algorithmic improvements from Li and Wood (2020). The examples above were all run on an Apple M1 Pro MacBook Pro with 32GB of RAM, and the entire analysis including the generations of figures takes only a few minutes.

GAMs are a very broad and general class of models, and many extensions exist in the literature. In the case of the pig growth data, several animals exhibited visibly more variation in their weight measurements than the majority of the animals in the data. Such heteroscedasticity (non-constant variance) can be modelled using distributional GAMs (or Generalized additive models for location, scale, shape or GAMLSS) (e.g. Rigby and Stasinopoulos, 2005; Kneib et al., 2021; Klein, 2024), which include linear predictors for all of the parameters of a distribution, or centile or quantile models (e.g., Nakamura et al., 2022). In the examples above, the additional parameters (scale in the case of the Gamma, or the Tweedie power parameter) were estimated as constants for the entire data set. A distributional GAM would allow those parameters to potentially vary with individual animals, or as smooth functions of the covariates, just as was done for the mean of the distribution in this study.

In conclusion, GAMs are a modern, flexible, and highly usable statistical model that is amenable to many research problems in animal science, and deserve a place in the

711 statistical toolbox.

#### 712 **Ethics approval**

713 Not applicable. See the original sources of the data used for ethical approval.

#### 714 **Declaration of Generative AI and AI-assisted technologies in the writing process**

715 The author did not use any artificial intelligence technologies in the writing process.

#### 716 **Author ORCIDs**

717 **Gavin L. Simpson:** 0000-0002-9084-8413

#### 718 **Author contributions**

719 GLS: Conceptualization, Methodology, Software, Formal analysis, Writing, Visualiza-  
720 tion.

#### 721 **Declaration of interest**

722 None.

#### 723 **Acknowledgements**

724 The author would to like to express their appreciation to colleagues at the Department  
725 of Animal and Veterinary Sciences, Aarhus University for making the pig growth data  
726 available, and in particular to Dr. Mona Larsen for supplying the data and arranging with  
727 their coauthors to enable the subset analysed in this manuscript to be made available  
728 as open data.

729 **Financial support statement**

730 This work was supported by an Aarhus Universitets Forskningsfond (Aarhus University  
731 Research Foundation; AUFF) starting grant awarded to the author.

732 **References**

- 733 Arel-Bundock, V., Greifer, N., Heiss, A., 2024. How to interpret statistical models using marginaeffects for  
734 *R* and *Python*. *J. Stat. Softw.* 111, 1–32. URL: <https://www.jstatsoft.org/index.php/jss/article/view/v111i09>,  
735 doi:10.18637/jss.v111.i09.
- 736 Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.P., Delignette-Muller, M.L., 2015. A toolbox for  
737 nonlinear regression in R: The package nlstools. *Journal of Statistical Software* 66, 1–21. doi:10.18637/  
738 jss.v066.i05.
- 739 Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency.  
740 *Ann. Stat.* 29, 1165–1188. URL: [http://projecteuclid.org/download/pdf\\_1/euclid.aos/1013699998](http://projecteuclid.org/download/pdf_1/euclid.aos/1013699998).
- 741 Benni, S., Pastell, M., Bonora, F., Tassinari, P., Torreggiani, D., 2020. A generalised additive model to  
742 characterise dairy cows' responses to heat stress. *Animal* 14, 418–424. URL: [http://dx.doi.org/10.1017/](http://dx.doi.org/10.1017/S1751731119001721)  
743 [S1751731119001721](http://dx.doi.org/10.1017/S1751731119001721), doi:10.1017/S1751731119001721.
- 744 Brito, L.F., Gomes da Silva, F., Rojas de Oliveira, H., Souza, N., Caetano, G., Costa, E.V., Romeiro de  
745 Oliveira Menezes, G., Puerro de Melo, A.L., Teixeira Rodrigues, M., de Almeida Torres, R., 2017. Modelling  
746 lactation curves of dairy goats by fitting random regression models using legendre polynomials or b-splines.  
747 *Can. J. Anim. Sci.* URL: <http://dx.doi.org/10.1139/cjas-2017-0019>, doi:10.1139/cjas-2017-0019.
- 748 Bus, J.D., Franchi, G.A., Boumans, I.J.M.M., Te Beest, D.E., Webb, L.E., Jensen, M.B., Pedersen, L.J.,  
749 Bokkers, E.A.M., 2025. Short-term associations between ambient ammonia concentrations and growing-  
750 finishing pig performance and health. *Prev. Vet. Med.* 242, 106555. URL: [http://dx.doi.org/10.1016/j.](http://dx.doi.org/10.1016/j.prevetmed.2025.106555)  
751 [prevetmed.2025.106555](http://dx.doi.org/10.1016/j.prevetmed.2025.106555), doi:10.1016/j.prevetmed.2025.106555.
- 752 Dijkstra, J., Lopez, S., Bannink, A., Dhanoa, M.S., Kebreab, E., Odongo, N.E., Fathi Nasri, M.H., Behera,  
753 U.K., Hernandez-Ferrer, D., France, J., 2010. Evaluation of a mechanistic lactation model using cow, goat  
754 and sheep data. *J. Agric. Sci.* 148, 249–262. URL: <http://dx.doi.org/10.1017/S0021859609990578>, doi:10.  
755 [1017/s0021859609990578](http://dx.doi.org/10.1017/S0021859609990578).
- 756 Eilers, P.H.C., Marx, B.D., 1996. Flexible smoothing with B-splines and penalties. *Stat. Sci.* 11, 89–  
757 121. URL: [https://projecteuclid.org/journals/statistical-science/volume-11/issue-2/Flexible-smoothing-with-](https://projecteuclid.org/journals/statistical-science/volume-11/issue-2/Flexible-smoothing-with-B-splines-and-penalties/10.1214/ss/1038425655.short)  
758 [B-splines-and-penalties/10.1214/ss/1038425655.short](https://projecteuclid.org/journals/statistical-science/volume-11/issue-2/Flexible-smoothing-with-B-splines-and-penalties/10.1214/ss/1038425655.short), doi:10.1214/ss/1038425655.
- 759 Franchi, G.A., Bus, J.D., Boumans, I.J.M.M., Bokkers, E.A.M., Jensen, M.B., Pedersen, L.J., 2023. Estimating  
760 body weight in conventional growing pigs using a depth camera. *Smart Agric. Technol.* 3, 100117. URL:  
761 <http://dx.doi.org/10.1016/j.atech.2022.100117>, doi:10.1016/j.atech.2022.100117.



- Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman & Hall / CRC. URL: <https://market.android.com/details?id=book-qa29r1Ze1coC>.
- Henderson, H.V., McCulloch, C., 1990. Transform or link. URL: <https://ecommons.cornell.edu/bitstream/1813/31620/1/BU-1049-MA.pdf>.
- Hirst, W.M., Murray, R.D., Ward, W.R., French, N.P., 2002. Generalised additive models and hierarchical logistic regression of lameness in dairy cows. *Prev. Vet. Med.* 55, 37–46. URL: [http://dx.doi.org/10.1016/S0167-5877\(02\)00058-2](http://dx.doi.org/10.1016/S0167-5877(02)00058-2), doi:10.1016/S0167-5877(02)00058-2.
- Huang, C.H., Furukawa, K., Kusaba, N., 2023. Estimating the nonlinear interaction between somatic cell score and differential somatic cell count on milk production by parity using generalized additive models. *J. Dairy Sci.* URL: <http://dx.doi.org/10.3168/jds.2022-22958>, doi:10.3168/jds.2022-22958.
- Klein, N., 2024. Distributional regression for data analysis. *Annu. Rev. Stat. Appl.* 11. URL: <https://www.annualreviews.org/content/journals/10.1146/annurev-statistics-040722-053607>, doi:10.1146/annurev-statistics-040722-053607.
- Kneib, T., Silbersdorff, A., Säfken, B., 2021. Rage against the mean – a review of distributional regression approaches. *Econometrics and Statistics* URL: <https://www.sciencedirect.com/science/article/pii/S2452306221000824>, doi:10.1016/j.ecosta.2021.07.006.
- Li, Z., Wood, S.N., 2020. Faster model matrix crossproducts for large generalized linear models with discretized covariates. *Stat. Comput.* 30, 19–25. URL: <https://doi.org/10.1007/s11222-019-09864-2>, doi:10.1007/s11222-019-09864-2.
- van Lingen, H.J., Fadel, J.G., Kebreab, E., Bannink, A., Dijkstra, J., van Gastelen, S., 2023. Smoothing spline assessment of the accuracy of enteric hydrogen and methane production measurements from dairy cattle using various sampling schemes. *J. Dairy Sci.* 106, 6834–6848. URL: <http://dx.doi.org/10.3168/jds.2022-23207>, doi:10.3168/jds.2022-23207.
- Macciotta, N.P.P., Dimauro, C., Rassu, S.P.G., Steri, R., Pulina, G., 2011. The mathematical description of lactation curves in dairy cattle. *Ital. J. Anim. Sci.* 10, e51. URL: <http://dx.doi.org/10.4081/ijas.2011.e51>, doi:10.4081/ijas.2011.e51.
- Macciotta, N.P.P., Miglior, F., Dimauro, C., Schaeffer, L.R., 2010. Comparison of parametric, orthogonal, and spline functions to model individual lactation curves for milk yield in canadian holsteins. *Ital. J. Anim. Sci.* 9, e87. URL: <https://doi.org/10.4081/ijas.2010.e87>, doi:10.4081/ijas.2010.e87.
- Miller, D.L., 2025. Bayesian views of generalized additive modelling. *Methods Ecol. Evol.* URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.14498>, doi:10.1111/2041-210X.14498.
- Morota, G., Cheng, H., Cook, D., Tanaka, E., 2021. Asas-nanp symposium: prospects for interactive and dynamic graphics in the era of data-rich animal science1. *J. Anim. Sci.* 99. URL: <http://dx.doi.org/10.1093/jas/skaa402>, doi:10.1093/jas/skaa402.
- Nagel-Alne, G.E., Krøntveit, R., Bohlin, J., Valle, P.S., Skjerve, E., Sølverød, L.S., 2014. The norwegian

healthier goats program—modeling lactation curves using a multilevel cubic spline regression model. J. Dairy Sci. 97, 4166–4173. URL: <http://dx.doi.org/10.3168/jds.2013-7228>, doi:10.3168/jds.2013-7228.

Nakamura, L.R., Ramires, T.G., Righetto, A.J., Pescim, R.R., Roquim, F.V., Savian, T.V., Stasinopoulos, D.M., 2022. Cattle reference growth curves based on centile estimation: A gamlss approach. Comput. Electron. Agric. 192, 106572. URL: <http://dx.doi.org/10.1016/j.compag.2021.106572>, doi:10.1016/j.compag.2021.106572.

Oliveira, J.G.d., Sant’Anna, D.F.D., Lourenço, M.C., Tavares, D.S.T., Rodrigues, M.T., Tedeschi, L.O., Vieira, R.A.M., 2020. The geometry of the lactation curve based on wood’s equation: a two-step prediction. Rev. Bras. Zootec. 49, —. URL: <http://dx.doi.org/10.37496/rbz4920200023>, doi:10.37496/rbz4920200023.

Pedersen, E.J., Miller, D.L., Simpson, G.L., Ross, N., 2019. Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ 7, e6876. URL: <http://dx.doi.org/10.7717/peerj.6876>, doi:10.7717/peerj.6876.

Pya, N., Wood, S.N., 2016. A note on basis dimension selection in generalized additive modelling. arXiv [stat.ME] URL: <http://arxiv.org/abs/1602.06696>, arXiv:1602.06696.

R Core Team, 2025. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.

Reiss, P.T., Ogden, R.T., 2009. Smoothing parameter selection for a class of semiparametric linear models. J. R. Stat. Soc. Series B Stat. Methodol. 71, 505–523. URL: <http://doi.wiley.com/10.1111/j.1467-9868.2008.00695.x>, doi:10.1111/j.1467-9868.2008.00695.x.

Rigby, R.A., Stasinopoulos, D.M., 2005. Generalized additive models for location, scale and shape. J. R. Stat. Soc. Ser. C Appl. Stat. 54, 507–554. URL: <http://dx.doi.org/10.1111/j.1467-9876.2005.00510.x>, doi:10.1111/j.1467-9876.2005.00510.x.

Sarraude, T., Hsu, B.Y., Groothuis, T., Ruuskanen, S., 2020a. Dataset of prenatal thyroid hormones manipulation in Japanese quails. URL: <https://zenodo.org/record/3741711>, doi:10.5281/zenodo.3741711.

Sarraude, T., Hsu, B.Y., Groothuis, T., Ruuskanen, S., 2020b. Testing the short-and long-term effects of elevated prenatal exposure to different forms of thyroid hormones. PeerJ 8, e10175. URL: <http://dx.doi.org/10.7717/peerj.10175>, doi:10.7717/peerj.10175.

Silvestre, A.M., Petim-Batista, F., Colaço, J., 2006. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. J. Dairy Sci. 89, 1813–1821. URL: <http://www.journalofdairyscience.org/article/S0022030206722500/abstract>, doi:10.3168/jds.S0022-0302(06)72250-0.

Simpson, G.L., 2024. gratia: An R package for exploring generalized additive models. J. Open Source Softw. 9, 6962. URL: <https://joss.theoj.org/papers/10.21105/joss.06962>, doi:10.21105/joss.06962.

Smyth, G., 1996. Regression modelling quantity data exact zeros, in: Wilson, R.J., Osaki, S., Murthy, D.N.P. (Eds.), Proceedings second Australia-Japan workshop stochastic models engineering, Technology Manage-

ment Centre, The University of Queensland Technology Management Centre, The University of Queensland.  
pp. 572–580.

White, I.M., Thompson, R., Brotherstone, S., 1999. Genetic and environmental smoothing of lactation curves with cubic splines. *J. Dairy Sci.* 82, 632–638. URL: [http://dx.doi.org/10.3168/jds.S0022-0302\(99\)75277-X](http://dx.doi.org/10.3168/jds.S0022-0302(99)75277-X), doi:10.3168/jds.S0022-0302(99)75277-X.

Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. Use R!, Springer International Publishing. URL: <https://link.springer.com/book/10.1007/978-3-319-24277-4>, doi:10.1007/978-3-319-24277-4.

Wood, P.D.P., 1967. Algebraic model of the lactation curve in cattle. *Nature* 216, 164–165. URL: <http://dx.doi.org/10.1038/216164a0>, doi:10.1038/216164a0.

Wood, S.N., 2003. Thin plate regression splines. *J. R. Stat. Soc. Series B Stat. Methodol.* 65, 95–114. URL: <http://dx.doi.org/10.1111/1467-9868.00374>, doi:10.1111/1467-9868.00374.

Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semi-parametric generalized linear models. *J. R. Stat. Soc. Series B Stat. Methodol.* 73, 3–36. URL: <http://dx.doi.org/10.1111/j.1467-9868.2010.00749.x>, doi:10.1111/j.1467-9868.2010.00749.x.

Wood, S.N., 2013a. On p-values for smooth components of an extended generalized additive model. *Biometrika* 100, 221–228. URL: <http://biomet.oxfordjournals.org/content/100/1/221.abstract>, doi:10.1093/biomet/ass048.

Wood, S.N., 2013b. A simple test for random effects in regression models. *Biometrika* 100, 1005–1010. URL: <http://biomet.oxfordjournals.org/content/100/4/1005.abstract>, doi:10.1093/biomet/ast038.

Wood, S.N., 2017. *Generalized Additive Models: An Introduction with R*, Second Edition. CRC Press. URL: <https://market.android.com/details?id=book-JTkkDwAAQBAJ>.

Wood, S.N., 2025. *mgcv: Mixed GAM computation vehicle with automatic smoothness estimation*. URL: <http://dx.doi.org/10.32614/cran.package.mgcv>, doi:10.32614/cran.package.mgcv.

Wood, S.N., Pya, N., Säfken, B., 2016. Smoothing parameter and model selection for general smooth models. *J. Am. Stat. Assoc.* 111, 1548–1563. URL: <https://doi.org/10.1080/01621459.2016.1180986>, doi:10.1080/01621459.2016.1180986, arXiv:<http://dx.doi.org/10.1080/01621459.2016.1180986>, doi: 10.1080/01621459.2016.1180986.

Yano, M., Shimadzu, H., Endo, T., 2014. Modelling temperature effects on milk production: a study on holstein cows at a japanese farm. *SpringerPlus* 3, 129. URL: <http://dx.doi.org/10.1186/2193-1801-3-129>, doi:10.1186/2193-1801-3-129.

Table 1: Comparison of models fitted to the lactation curve data set. Degrees of freedom and effective degrees of freedom ((E)DF) are shown for the Wood model and Tweedie GLM, and for the Tweedie GAM, respectively, showing the complexity in terms of (effective) numbers of parameters in each model. Akaike's An information criterion (AIC) values are shown for each model, alongside an estimate of the root mean squared error of the model fit estimated from the response residuals of each model.

Model	(E)DF	AIC	RMSE
Wood	4	-100.61	0.30
Tweedie GLM	5	-82.76	0.32
Tweedie GAM	8.004	-132.97	0.15

Table 2: Comparison of models fitted to the pig weight data set. The model label is shown (see text for the specific model formulations). Effective degrees of freedom (EDF) represents model complexity in terms of the (effective) number of parameters in each model. Akaike's An information criterion (AIC) values, model deviance, and deviance explained as a proportion are also reported.

Model	EDF	AIC	Deviance	Deviance expl.
P1	83.542	3838.987	2.074	0.964
P2	64.386	3797.407	2.088	0.964
P3	90.570	3823.754	2.050	0.964
P4	64.754	3809.342	2.148	0.963

Table 3: Summary for model P2 fitted to the pig growth data set. The model term for comparison with the descriptions in the text and the label reported by the software are both shown for clarity.  $k$  is the number of basis functions *per smooth*, EDF is the effective degrees of freedom, a measure of the complexity of each term,  $\text{EDF}_{\text{Ref.}}$  is the reference degrees of freedom used in the tests, F is the test statistic and  $p$  the  $p$  value of the null hypothesis of a flat constant function or functions. The tests for terms in rows 1–3 are Wald-like tests with appropriate degrees of freedom. The tests of smooth terms (rows 2–3) are described in [Wood \(2013a\)](#).

Model term	Label	$k$	EDF	$\text{EDF}_{\text{Ref.}}$	F	$p$
$\beta_0$	Intercept	NA	1.000	1.000	1328.642	<0.001
$f(\text{day})$	s(day)	9	7.332	8.124	1302.500	<0.001
$f_{a(i)}(\text{day})$	s(day,animal)	10	56.054	63.309	10.033	<0.001

Table 4: Comparison of models fitted to the quail hormone experimental data set. The model label is shown (see text for the specific model formulations). Effective degrees of freedom (EDF) represents model complexity in terms of the (effective) number of parameters in each model. Akaike's An information criterion (AIC) values, model deviance, and deviance explained as a proportion are also reported.

Model	EDF	AIC	Deviance	Deviance expl.
Q1	275.028	4458.053	12.229	0.998
Q2	274.839	4458.919	12.360	0.998
Q3	274.755	4459.222	12.528	0.998
Q4	77.115	4986.610	30.416	0.995

Table 5: Summary for model Q3 fitted to the quail hormone experimental data set. The model term for comparison with the descriptions in the text and the label reported by the software are both shown for clarity.  $k$  is the number of basis functions *per smooth*, EDF is the effective degrees of freedom, a measure of the complexity of each term,  $\text{EDF}_{\text{Ref.}}$  is the reference degrees of freedom used in the tests, F is the test statistic and  $p$  the  $p$  value of the null hypothesis of a flat constant function or functions. The tests for terms in rows 1–6 are Wald-like tests with appropriate degrees of freedom. The tests of smooth terms (rows 2–6) are described in [Wood \(2013a\)](#). The test of the random intercept term (row 7) is adjusted for testing a variance component whose null hypothesis lies on the boundary of the allowable parameter space ([Wood, 2013b](#)).

Model term	Label	$k$	EDF	$\text{EDF}_{\text{Ref.}}$	F	$p$
$\beta_0$	Intercept	NA	1.000	1.000	142.090	<0.001
$f(\text{day})$	s(day)	9	8.947	8.987	4797.205	<0.001
$f_{\text{treat}(i)}(\text{day}_i)$	s(day,treat)	10	6.007	6.009	1.658	0.129
$f_{\text{sex}(i)}(\text{day}_i)$	s(day,sex)	10	7.318	8.249	14.312	<0.001
$f_{\text{treat}(i),\text{sex}(i)}(\text{day}_i)$	s(day,treat,sex)	10	9.729	10.712	1.561	0.111
$f_{\text{egg}(i)}(\text{day}_i)$	s(day,egg)	6	224.064	326.000	67.697	0.267
$\xi_{\text{mother}(i)}$	s(mother)	21	17.690	21.000	8.633	<0.001



Table 6: Pairwise comparisons of treatment effects by sex on the slopes of the growth curves at day 20 for an average quail of the indicated sex in the pair of treatments shown. The Hypothesis column lists the specific pairwise comparison, Diff. is the estimated difference in the slope of the growth curves compared, SE its standard error. Z is the Wald test statistic,  $p$  its  $p$  value, and associated endpoints of a Bayesian 95% credible interval on the estimated difference of slopes.

Sex	Hypothesis	Diff.	SE	Z	$p$	2.5%	97.5%
Female	$T_4 - \text{Control}$	-0.399	0.276	-1.445	1	-1.171	0.373
Female	$T_3 - \text{Control}$	-0.287	0.262	-1.097	1	-1.019	0.445
Female	$T_3 T_4 - \text{Control}$	-0.341	0.269	-1.266	1	-1.094	0.413
Female	$T_3 - T_4$	0.112	0.207	0.541	1	-0.467	0.691
Female	$T_3 T_4 - T_4$	0.058	0.257	0.225	1	-0.660	0.776
Female	$T_3 T_4 - T_3$	-0.054	0.218	-0.248	1	-0.664	0.556
Male	$T_4 - \text{Control}$	-0.288	0.251	-1.148	1	-0.992	0.415
Male	$T_3 - \text{Control}$	-0.144	0.273	-0.529	1	-0.907	0.619
Male	$T_3 T_4 - \text{Control}$	-0.203	0.233	-0.871	1	-0.856	0.450
Male	$T_3 - T_4$	0.144	0.252	0.572	1	-0.560	0.849
Male	$T_3 T_4 - T_4$	0.085	0.274	0.310	1	-0.682	0.852
Male	$T_3 T_4 - T_3$	-0.059	0.261	-0.226	1	-0.790	0.672

Table 7: Pairwise comparisons of treatment effects by sex on the estimated mean of the quail growth curves at day 78 for an average quail of the indicated sex in each pair of treatments. The Hypothesis column lists the specific pairwise comparison, Diff. is the estimated difference of mean body mass (g) of the growth curves compared, SE its standard error.  $Z$  is the Wald test statistic,  $p$  its  $p$  value, and associated endpoints of a Bayesian 95% credible interval on the estimated difference of means.

Sex	Hypothesis	Diff.	SE	$Z$	$p$	2.5%	97.5%
Female	$T_4 - \text{Control}$	-9.985	5.705	-1.750	1	-25.958	5.988
Female	$T_3 - \text{Control}$	-6.279	5.817	-1.079	1	-22.566	10.007
Female	$T_3 T_4 - \text{Control}$	-5.529	5.589	-0.989	1	-21.178	10.119
Female	$T_3 - T_4$	3.706	4.021	0.922	1	-7.552	14.964
Female	$T_3 T_4 - T_4$	4.456	3.633	1.227	1	-5.715	14.626
Female	$T_3 T_4 - T_3$	0.750	4.341	0.173	1	-11.405	12.904
Male	$T_4 - \text{Control}$	-0.684	4.501	-0.152	1	-13.285	11.917
Male	$T_3 - \text{Control}$	2.558	5.374	0.476	1	-12.488	17.604
Male	$T_3 T_4 - \text{Control}$	-2.921	4.017	-0.727	1	-14.168	8.325
Male	$T_3 - T_4$	3.242	4.684	0.692	1	-9.873	16.356
Male	$T_3 T_4 - T_4$	-2.238	3.705	-0.604	1	-12.612	8.136
Male	$T_3 T_4 - T_3$	-5.479	4.808	-1.140	1	-18.941	7.982

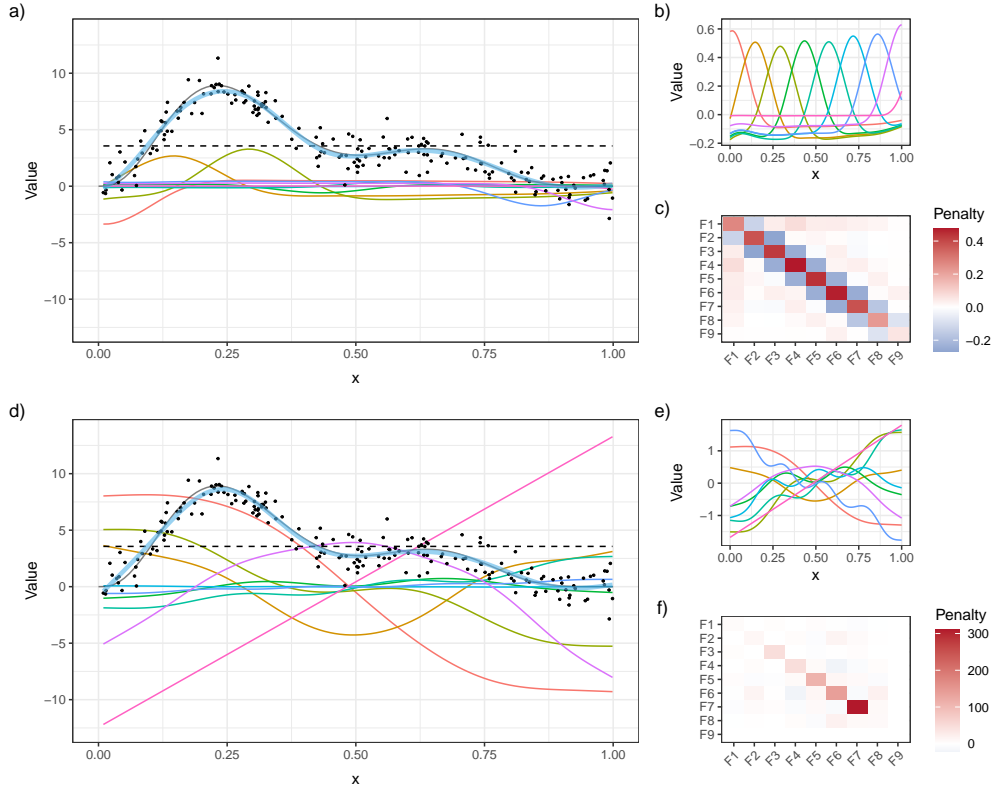


Fig. 1: Illustration of how penalised splines work. A spline basis expansion (a) and associated penalty matrix  $S$  (c) are formed for a covariate  $x$ . Model fitting involves finding estimates for the coefficients of the basis functions that make the fitted spline (thick, blue curve) go as close to the data (black points) as possible, without over fitting (a). In (a) and (b) the basis functions are shown as thin coloured lines and are from a cubic B spline basis. The thin black line in (a) is the true function from which the data were simulated. The dashed horizontal line in (a) is the estimated value of the intercept. The sum-to-zero identifiability constraint needed so that an intercept can be included in the model has been absorbed into the basis shown in (a) and (b). The penalty matrix (c) encodes how wiggly each basis function is in terms of its second derivative. Panels (d) – (e) are as above but for a low-rank thin plate spline basis.

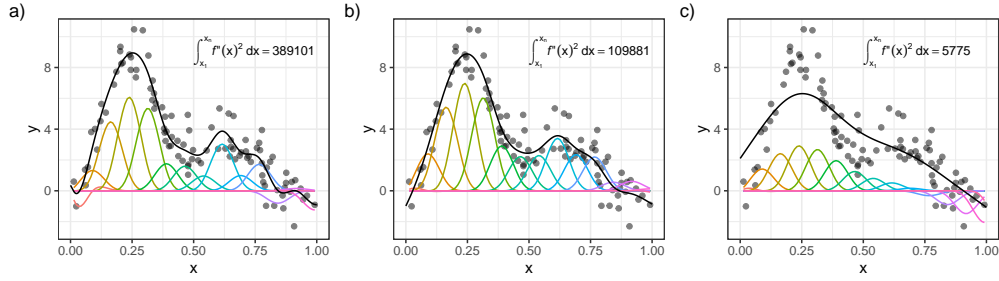


Fig. 2: Illustration of how the wigginess penalty controls the resulting fit of a penalised spline. The weighted basis functions are shown as thin coloured lines. In each panel a penalised spline is shown by the solid black line, which has been fitted to the data points shown. The wigginess value of the spline, the integrated squared derivative of the fitted spline over  $x$  is given in the upper right of each panel. The spline in (a) is over fitted to the data, resulting in a very wiggly function with a large wigginess value. The spline in (c) is over smoothed, resulting in a simple fitted function with low wigginess, but which does not fit the data well. The spline in (b) represents a balance between fit to the data and complexity of fitted function. The smoothing parameter for the spline,  $\lambda$ , is used as a tuning parameter in the model, which ultimately controls this balance between fit and complexity.

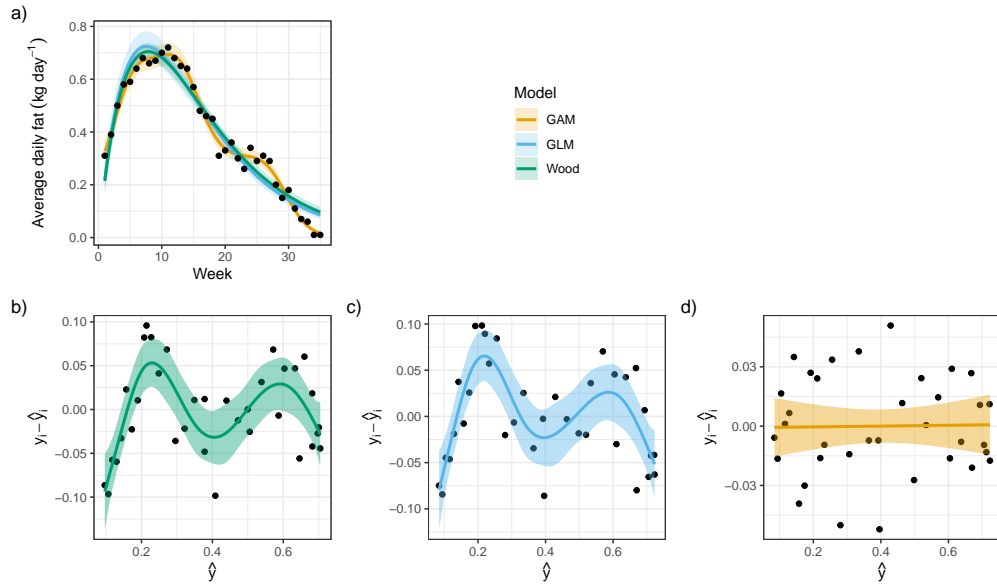


Fig. 3: Results of model fitting to the average daily fat content data from [Henderson and McCulloch \(1990\)](#). a) observed average daily fat content (points) and estimated lactation curves from Wood's (1967) model, a Tweedie GLM, and a Tweedie GAM (lines) with associated 95% confidence (Wood's model) or 95% credible intervals (GLM and GAM). Response residuals for Wood's model (b), Tweedie GLM (c), and Tweedie GAM (d), plus scatter plot smoothers (lines) and 95% credible intervals (shaded ribbons).

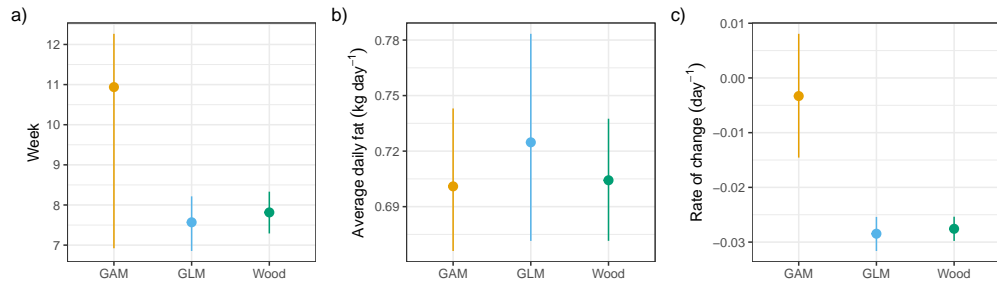


Fig. 4: Quantities of interest derived from Wood's model, a Tweedie GLM, and a Tweedie GAM fitted to the lactation data example: a) the estimated week of peak average daily fat content, b) the estimated average daily fat content at the peak, and c) the rate of change (first derivative) of the lactation curve estimated at a point that is midway between the peak fat content and the end of lactation. The points are the estimated values and the lines are a 95% uncertainty interval. The uncertainty interval is based on the 0.025 and 0.975 percentiles of the bootstrap distribution of model coefficient estimates (Wood's model) or of the posterior distribution (GLM and GAM).

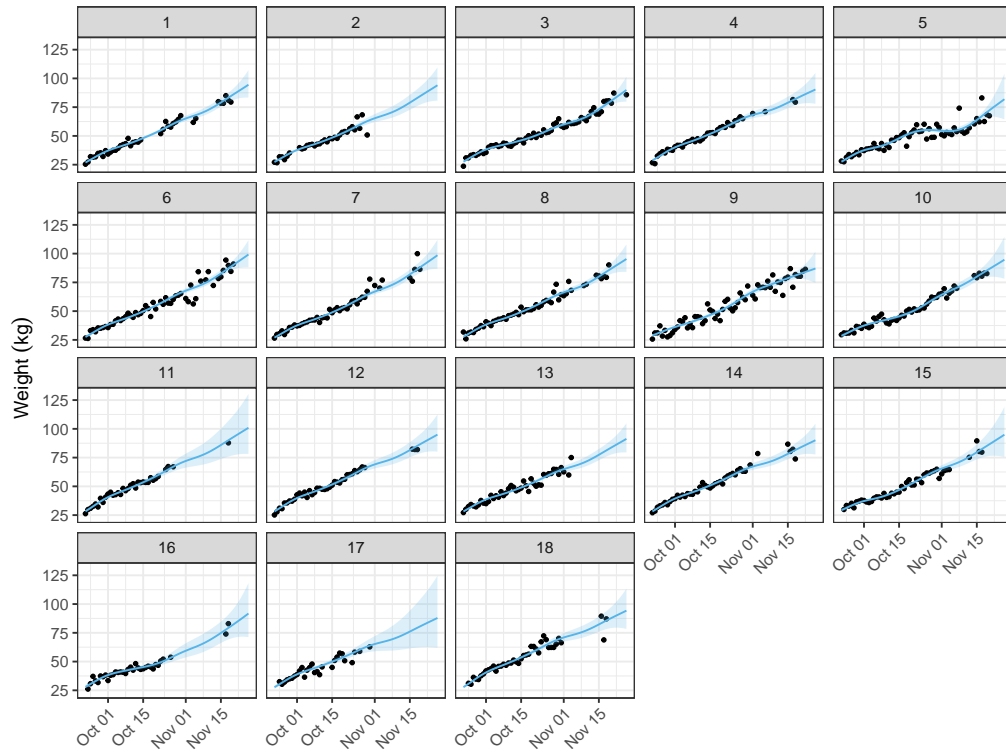


Fig. 5: Depth camera-based weight estimates from 18 commercial pigs. The data (black points) are the average of multiple measurements taken of each animal per day, 1 panel per pig. The panel labels indicate the pig shown. The estimated growth curve for each pig obtained using generalized additive model P2 is shown by the blue line in each panel. The blue shaded ribbon is the 95% Bayesian credible interval around the estimate curve.

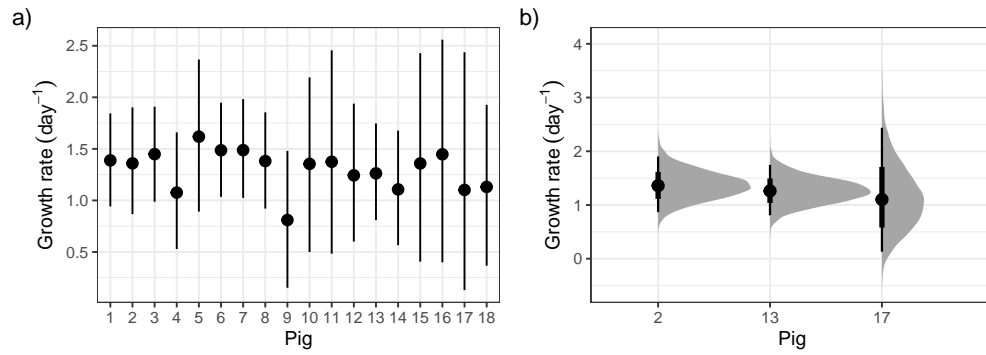


Fig. 6: a) Estimated daily growth rate on November 15<sup>th</sup>, 2021 and 95% Bayesian credible interval for the 18 pigs in the pig growth example. b) Posterior distribution of daily growth rate on November 15<sup>th</sup>, 2021, for three pigs (numbers 2, 13, and 17), for whom weight observations ceased before November 1<sup>st</sup>, 2021. In b), the shaded region is the posterior distribution, the point, and thick and thin bars are the posterior median, and 66% and 95% posterior intervals respectively.



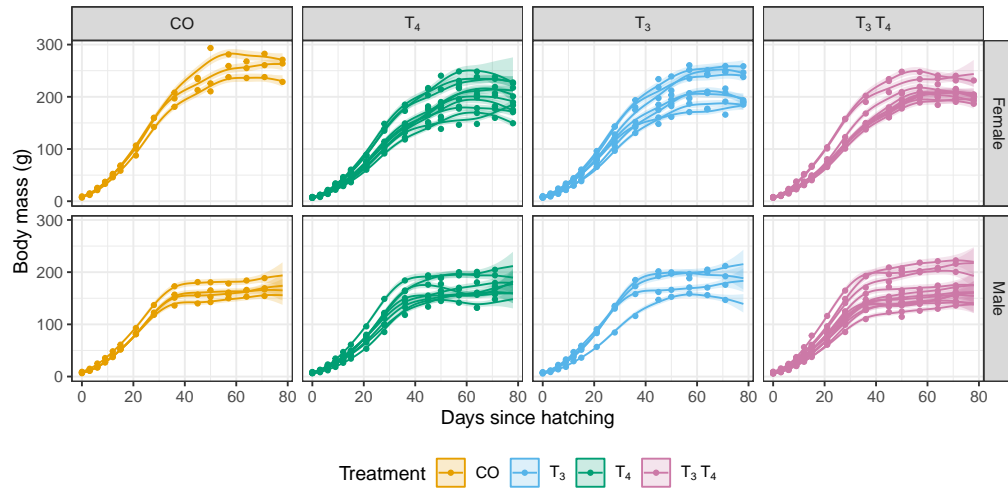


Fig. 7: Measured body mass (g) for 57 Japanese quail from an experiment on the effects of exposure to the maternal hormone T<sub>4</sub>, its active metabolite T<sub>3</sub>, both (T<sub>3</sub>T<sub>4</sub>) or a saline solution control (CO). The data (points) are shown along side the estimated growth curve for each quail obtained using generalized additive model Q3, which are shown by the coloured lines in each panel. The coloured shaded ribbon is the 95% Bayesian credible interval around the estimate curve. The data are faceted by treatment and the sex of bird.

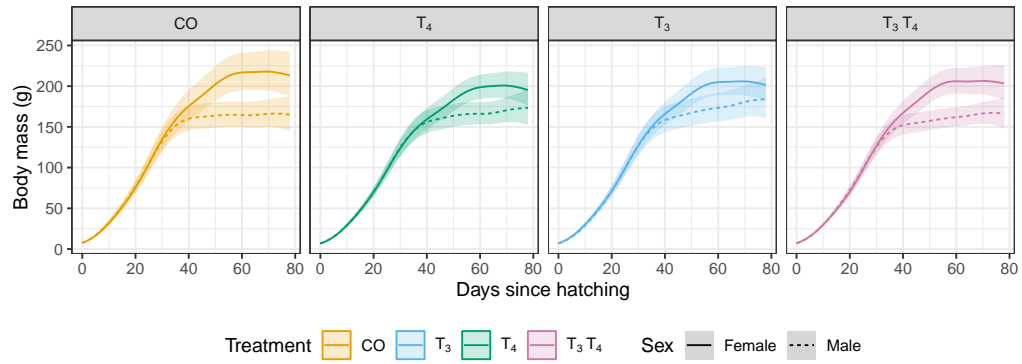


Fig. 8: Conditional value plots showing the growth rate for an average quail in each of the treatment groups by sex. The panels show the estimated curves for a particular treatment group; saline solution controls (CO), maternal hormone metabolite  $T_3$ , the maternal hormone  $T_4$ , and a combination of both  $T_3T_4$ . The estimated curve for male birds is shown by the solid line, and females the dashed line in each panel. The shaded band around each curve is a 95% Bayesian credible interval.

## 862 Figure captions

863 Fig. 1: Illustration of how penalised splines work. A spline basis expansion (a) and  
 864 associated penalty matrix  $S$  (c) are formed for a covariate  $x$ . Model fitting involves finding  
 865 estimates for the coefficients of the basis functions that make the fitted spline (thick, blue  
 866 curve) go as close to the data (black points) as possible, without over fitting (a). In (a)  
 867 and (b) the basis functions are shown as thin coloured lines and are from a cubic B spline  
 868 basis. The thin black line in (a) is the true function from which the data were simulated.  
 869 The dashed horizontal line in (a) is the estimated value of the intercept. The sum-to-  
 870 zero identifiability constraint needed so that an intercept can be included in the model  
 871 has been absorbed into the basis shown in (a) and (b). The penalty matrix (c) encodes  
 872 how wiggly each basis function is in terms of its second derivative. Panels (d) – (e) are  
 873 as above but for a low-rank thin plate spline basis.

874 Fig. 2: Illustration of how the wiggleness penalty controls the resulting fit of a penalised  
 875 spline. The weighted basis functions are shown as thin coloured lines. In each panel a  
 876 penalised spline is shown by the solid black line, which has been fitted to the data points  
 877 shown. The wiggleness value of the spline, the integrated squared derivative of the fitted  
 878 spline over  $x$  is given in the upper right of each panel. The spline in (a) is over fitted to  
 879 the data, resulting in a very wiggly function with a large wiggleness value. The spline in  
 880 (c) is over smoothed, resulting in a simple fitted function with low wiggleness, but which  
 881 does not fit the data well. The spline in (b) represents a balance between fit to the data  
 882 and complexity of fitted function. The smoothing parameter for the spline,  $\lambda$ , is used as  
 883 a tuning parameter in the model, which ultimately controls this balance between fit and  
 884 complexity.

885 Fig. 3: Results of model fitting to the average daily fat content data from [Henderson and](#)  
 886 [McCulloch \(1990\)](#). a) observed average daily fat content (points) and estimated lactation  
 887 curves from Wood's ([1967](#)) model, a Tweedie GLM, and a Tweedie GAM (lines) with  
 888 associated 95% confidence (Wood's model) or 95% credible intervals (GLM and GAM).

889 Response residuals for Wood's model (b), Tweedie GLM (c), and Tweedie GAM (d), plus  
890 scatter plot smoothers (lines) and 95% credible intervals (shaded ribbons).

891 Fig. 4: Quantities of interest derived from Wood's model, a Tweedie GLM, and a  
892 Tweedie GAM fitted to the lactation data example: a) the estimated week of peak av-  
893 erage daily fat content, b) the estimated average daily fat content at the peak, and c) the  
894 rate of change (first derivative) of the lactation curve estimated at a point that is midway  
895 between the peak fat content and the end of lactation. The points are the estimated val-  
896 ues and the lines are a 95% uncertainty interval. The uncertainty interval is based on the  
897 0.025 and 0.975 percentiles of the bootstrap distribution of model coefficient estimates  
898 (Wood's model) or of the posterior distribution (GLM and GAM).

899 Fig. 5: Depth camera-based weight estimates from 18 commercial pigs. The data  
900 (black points) are the average of multiple measurements taken of each animal per day,  
901 1 panel per pig. The panel labels indicate the pig shown. The estimated growth curve  
902 for each pig obtained using generalized additive model P2 is shown by the blue line in  
903 each panel. The blue shaded ribbon is the 95% Bayesian credible interval around the  
904 estimate curve.

905 Fig. 6: a) Estimated daily growth rate on November 15<sup>th</sup>, 2021 and 95% Bayesian  
906 credible interval for the 18 pigs in the pig growth example. b) Posterior distribution of  
907 daily growth rate on November 15<sup>th</sup>, 2021, for three pigs (numbers 2, 13, and 17), for  
908 whom weight observations ceased before November 1<sup>st</sup>, 2021. In b), the shaded region  
909 is the posterior distribution, the point, and thick and thin bars are the posterior median,  
910 and 66% and 95% posterior intervals respectively.

911 Fig. 7: Measured body mass (g) for 57 Japanese quail from an experiment on the  
912 effects of exposure to the maternal hormone  $T_4$ , its active metabolite  $T_3$ , both ( $T_3T_4$ )  
913 or a saline solution control (CO). The data (points) are shown along side the estimated  
914 growth curve for each quail obtained using generalized additive model Q3, which are

915 shown by the coloured lines in each panel. The coloured shaded ribbon is the 95%  
916 Bayesian credible interval around the estimate curve. The data are faceted by treatment  
917 and the sex of bird.

918 Fig. 8: Conditional value plots showing the growth rate for an average quail in each  
919 of the treatment groups by sex. The panels show the estimated curves for a particular  
920 treatment group; saline solution controls (CO), maternal hormone metabolite  $T_3$ , the  
921 maternal hormone  $T_4$ , and a combination of both  $T_3T_4$ . The estimated curve for male  
922 birds is shown by the solid line, and females the dashed line in each panel. The shaded  
923 band around each curve is a 95% Bayesian credible interval.