

Method: Using generalized additive models in the animal sciences

G. L. Simpson^a

^aAarhus 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 spline or polynomial terms. If it is desirable to learn the shape of the relationship from the data directly, 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 wiggliness 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 without producing an overly complex function. 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

formal statistical inference in a designed experiment of the effects of exposure to maternal hormones on subsequent growth in Japanese quail. 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 shown that 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 statistical toolbox.

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 frequently encountered in animal science research. Generalized additive models and automatic smoothness selection via penalised splines provide an attractive, flexible, data-driven statistical model that is capable of estimating these relationships. I provide a description of the generalized additive model and demonstrate its use on three typical data examples; i) a lactation curve, ii) growth curves in commercial pigs, and iii) a experiment on the effects of maternal hormones on growth rates in Japanese quail.

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

15 **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: https://doi.org/10.5281/zenodo.15777270

16

Introduction

Many research questions in the animal sciences involve nonlinear relationships between covariates and a response variable of interest. A classic example, with a long history of statistically-based and mathematically-based research, is the plethora of models that have been described for the estimation of lactation curves from test day data or automatic milking machines (e.g., [Macciotta et al., 2011](#)). Another, is the estimation of growth curves in the context of breeding and genetics (e.g., [White et al., 1999](#)). Despite the frequency with which such nonlinear relationships are encountered, surprisingly little use of generalized additive models (GAMs) has been seen in animal science to date. Notable exceptions include [Hirst et al. \(2002\)](#), [Yano et al. \(2014\)](#), [Huang et al. \(2023\)](#), and [Benni et al. \(2020\)](#).

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

More attention has been paid to the use of splines, frequently in comparisons against some form of polynomial-based model, especially Legendre polynomials (e.g., [Nagel-Alne et al., 2014](#); [Silvestre et al., 2006](#); [Macciotta et al., 2010](#); [Brito et al., 2017](#)). There, the focus has largely been on the choice of the number of knots in the spline basis expansion and on the placement of those knots. Modern GAMs largely make such choices redundant; with penalised splines, a wiggleness penalty is used to avoid over-fitting, and low-rank eigen bases, such as the low-rank thin plate regression spline basis of [Wood](#)

44 (2003) avoid the knot-placement issue for most problems.

45 A guide to the use of GAMs in the animal science setting, written with users in mind,
46 is needed to raise awareness of the utility of these flexible models and to promote best
47 practice in fitting GAMs. Below, I describe GAMs and demonstrate visually how they
48 work. Then I apply GAMs to three different examples representing typical data encoun-
49 tered in animal science. The examples below are supported by tutorials containing the
50 computer code needed to fit the models in the R statistical software (R Core Team, 2025),
51 which are available along side the source code for the manuscript itself.

52 **Materials and methods**

53 *Generalized additive models*

54 A basic GAM (Hastie and Tibshirani, 1990) has the following form

$$\begin{aligned} 55 \quad y_i &\sim \mathcal{D}(\mu_i, \phi) \\ 56 \quad \mathbb{E}(y_i) &= \mu_i \\ 57 \quad g(\mu_i) &= \mathbf{X}_i \boldsymbol{\gamma} + \sum_j f_j(\bullet), \quad i = 1, \dots, n; \quad j = 1, \dots, J \end{aligned}$$

58 where y_i is a univariate response variable of interest that is modelled on the scale
59 of a link function $g()$, $\mathcal{D}(\mu_i, \phi)$ is a distribution, typically from, though not limited to, the
60 exponential family of distributions, with mean μ_i and scale parameter ϕ . \mathbf{X}_i is the i th row
61 of the model matrix of any parametric terms (including the model intercept or constant
62 term), and $\boldsymbol{\gamma}$ the associated regression parameters. The f_j are J smooth functions of one
63 or more covariates; I use \bullet as a placeholder for this definition, but in the simplest case
64 of a smooth of a single continuous covariate we have $f_j(x_{ij})$, where x_{ij} is a univariate
65 covariate.

66 In the remainder of this section, I aim to present *just enough* detail about GAMs to af-
 67 ford the reader a general understanding of what is involved in estimating such a model
 68 so that they can appreciate how GAMs work, and how we aim to avoid overfitting or
 69 having to choose how complex each smooth function should be. The original approach
 70 (Hastie and Tibshirani, 1990) for fitting GAMs, known as *backfitting*, required the ana-
 71 lyst to specify how many degrees of freedom each function in the model should take *a*
 72 *priori*, which was viewed by many as being too subjective for more than exploratory anal-
 73 ysis. With modern automatic smoothness selection methods, this problem has largely
 74 been resolved, through the use of penalised splines and fast algorithms for smoothness
 75 selection.

76 *Penalised splines*

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

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

80 For identifiability reasons, the basis is subject to a sum-to-zero constraint to remove the
 81 constant function from the span of the basis, which allows a separate constant term
 82 or intercept in the model; this is desirable if we also want to include categorical (factor)
 83 terms in the model for example. Figure 1a shows a B spline basis ($K = 16$) for a covariate
 84 x after the absorbing the sum-to-zero constraint into the basis.

85 Fitting a spline then involves finding estimates for the coefficients of basis functions, β_k .
 86 These coefficients weight the individual basis functions as shown in Figure 1b. To find
 87 the value of the fitted spline at each value of x , we sum up the values of the weighted
 88 basis functions evaluated at each value of x . This yields the light blue curve in Figure 1b.
 89 The coefficients for the basis functions, β_k , are determined by forcing the fitted function

90 to go as close to the data as possible. As shown in Figure 1b, we largely recover the
91 true function from which the data were simulated.

92 Expanding x_j into many basis functions in this way means we must be cognizant of the
93 risks of over fitting the sample of data we have to hand. Taken to the extreme, we could
94 obtain an arbitrarily close fit to the data by using as many basis functions as there are
95 data (i.e., $K = n$), but all this would achieve in practice is the replacement of the data
96 with a set of coefficients.

97 This raises the question of how many basis functions should be used? One option, if
98 there are n unique values of the covariate x , is to use n basis functions, leading to a
99 situation where our model would have as many coefficients as data. However, we do
100 not gain anything by using n basis functions from a statistical viewpoint. In practice we
101 typically use $\mathcal{N} \ll n$ basis functions. Yet, even with just \mathcal{N} basis functions, we face the
102 very real situation that our model may overfit the data if \mathcal{N} is too large. To address this,
103 modern approaches to fitting GAMs use penalised splines.

104 If our aim is to avoid overfitting, we need to penalise highly complex fitted functions,
105 and hence need to define what we mean by “complex”. A complex function would be
106 one that “wiggles” about markedly as we move from low to high values of the covariate,
107 x . Such wiggling around implies that the function has a high amount of curvature, which
108 we measure using the second derivative of the fitted function. While there are other def-
109 initions for complexity that we use, this the typical measure used for splines in statistical
110 models. Therefore, to measure the wiggleness of a fitted function we need to integrate,
111 or sum up, the second derivative of the function f over the range of x

$$\int_x f''(x)^2 dx = \beta^T \mathbf{S} \beta \quad (1)$$

112 where f'' indicates the second derivative of f , and note that we are integrating the *square*
113 of the second derivative because we need to allow for both negative and positive curva-

ture as the function wiggles. Conveniently, we can compute this integral as a function of the model coefficients β as shown on the right hand side of Equation 1. S is a known penalty matrix, which encodes the complexity of the basis functions. The penalty matrix for the basis expansion shown in Figure 1a is displayed in Figure 1c.

Figure 2 shows three different estimated smooths for the simulated shown in Figure 1 and 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 relationships (Figure 2b).

We find estimates of the basis function coefficients, β_k , that make the fitted spline go as close to the data as possible; in practice, this is done by maximising the penalised log likelihood

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

where $\ell_p(\beta)$ is the penalised log likelihood and $\ell(\beta)$ the log likelihood of the data, given the β , 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, γ , including the intercept have been absorbed into β for convenience. The λ_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 λ_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 (Wood, 2003). The spline

138 basis used in the above description was a B spline basis, and for that basis the knots
 139 were placed at evenly-spaced quantiles of the covariate. In general, the exact position-
 140 ing of the knots is unimportant, so long as they are spread out over the range of the
 141 covariate. If we wish to avoid the knot placement altogether, we could replace the B
 142 spline basis with a low rank thin plate regression spline basis (TPRS). We start with a
 143 radial basis function at each unique value of the covariate. This full, rich basis is then
 144 transformed and truncated through an eigen decomposition to retain the k eigenvectors
 145 with the smallest eigenvalues. The decomposition removes much of the excessive wig-
 146 gliness that n basis functions would provide, while retaining many of the good properties
 147 of the original basis (Wood, 2003). The main downside of the TPRS basis is that it is
 148 computationally expensive to form the basis when setting up the model; for larger data
 149 problems, with many thousands of data, simpler bases such as the cubic regression
 150 spline or B spline may be preferred. In the *mgcv* software used here, the TPRS basis is
 151 the default basis used for univariate smooths (Wood, 2025, Wood (2011)).

152 The algorithms used to fit GAMs need to estimate the model coefficients, β , and choose
 153 appropriate values of the smoothing parameters, λ_j . This process is known as smooth-
 154 ness selection, and there are several approaches to smoothness selection that can be
 155 taken. One is treat the problem as one of prediction, and choose λ_j in such a way as
 156 to minimise the cross-validated prediction error of the model. In practice it would be
 157 computationally costly to actually cross-validate the model for fitting, so we approximate
 158 the prediction error by minimising the generalised cross validation (GCV) error, AIC, or
 159 similar measure. An alternative means of smoothness selection is to take a Bayesian
 160 view of the smoothing process (see Miller, 2025, for an accessible introduction to this
 161 viewpoint); in doing so, the $\beta_j^T \mathbf{S}_j \beta_j$ should be viewed as multivariate normal priors on
 162 the β . From this Bayesian view of smoothing, we find that the criterion we wish to min-
 163 imise is that of a mixed effects model. Hence, we can think of the wiggly parts of the f_j
 164 as being fancy random effects, while the smooth parts of the f_j are fixed effects, and
 165 the λ_j are inversely proportional to the random effect variances one would observe if

the model were fitted as a mixed effects model. The Bayesian approach to smoothing can involve fully bayesian estimation using Markov chain Monte Carlo (MCMC) or simulation free estimation via the integrated nested Laplace approximation (INLA), or, using the equivalence of splines and random effects, we can take an empirical bayesian approach, which yields the posterior modes of the β_j , or the maximum a posteriori (MAP) estimates.

Examples

In the remainder of the section I describe three representative examples that demonstrate the utility of GAMs to address problems in animal science.

Lactation curves

As a simple illustration of the benefits of GAMs to learn the functional form of a relationship between a response variable and a covariate, rather than impose one through a parametric model, I reanalyse a small data set of average daily fat content per week of milk from a single cow (Figure 3). The data were reported in [Henderson and McCulloch \(1990\)](#). Initially, I followed their ([Henderson and McCulloch, 1990](#)) analysis and fitted a Gamma generalized linear model (GLM) with log link function. However, subsequent analysis of this model (and the GAM alternative described below) showed that the average daily fat data are under dispersed relative to the assumed conditional distribution. Instead, a Tweedie GLM (log link) was fitted, which has the same response shape as the gamma GLM but is more easily compared with the Tweedie GAM described below. The Tweedie GLM fitted had the following form

$$\begin{aligned} \text{fat}_i &\sim \mathcal{T}(\mu_i, \phi) \\ \log(\mu_i) &= \beta_0 + \beta_1 \log(\text{week}_i) + \beta_2 \text{week}_i . \end{aligned}$$

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

$$191 \quad \text{fat}_i = \alpha \text{week}_i^\delta \exp(\kappa \text{week}_i) + \varepsilon$$

192 where α , δ , and κ are parameters whose values are to be estimated, and ε is a Gaus-
 193 sian error term. The linear predictors in the GLM and Wood's model are equivalent
 194 because of the log link function; $\beta_0 = \log \alpha$, $\beta_1 = \delta$, and $\beta_2 = \kappa$. However, we would not
 195 expect both models to produce the same fitted lactation curve as different distributional
 196 assumptions are being made; in the GLM version we assume the average daily fat val-
 197 ues are conditionally Tweedie distributed, while in Wood's model they are assumed to
 198 be conditionally Gaussian distributed.

199 These models were compared with a GAM version of the log-link, Tweedie GLM, where
 200 the fixed functional form for the lactation curve has been replaced by a smooth function
 201 to be estimated from the data

$$202 \quad \text{fat}_i \sim \mathcal{T}(\mu_i, \phi)$$

$$203 \quad \log(\mu_i) = \beta_0 + f(\text{week}_i)$$

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

209 *Pig growth*

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

222 Automated estimation of body weights is a useful on-farm method for continuously
223 monitoring growth of commercial pigs. I analyse a subset of the weight data reported by
224 [Franchi et al. \(2023\)](#) from a study by [Bus et al. \(2025\)](#), using data from a single pen of 18
225 pigs (Figure 4). The body weight data were obtained using a depth camera (iDOL65, dol-
226 sensors A/S, Aarhus, Denmark), and the each weight observation is the daily average of
227 multiple measurements made by the camera. As the number of weight measurements
228 made each day varied per animal and per day, each weight observation in the data is the
229 average of a variable number of measurements. To accommodate this, the model included
230 the number of measurements averaged as an observation weight, where the precision
231 of each response data is proportional to the number of measurements averaged.

232 The data exhibit common and animal-specific variation. To model these features, four
233 different GAMs were fitted to the pig weight data. All of the models ultimately provide
234 animal-specific estimates of weight over time, but they decompose the growth curves in
235 different ways. The weight data were assumed to be conditionally gamma distributed,

with log link, $\text{weight}_i \sim \mathcal{G}(\mu_i, \phi)$, with $\log(\mu_i) = \eta_i$, where η is the linear predictor. The linear predictors for each of the four models were:

$$\begin{aligned} \text{P1: } \eta_i &= \beta_{a(i)} + f_{a(i)}(\text{day}_i) \\ \text{P2: } \eta_i &= \beta_0 + f_1(\text{day}_i) + f_{a(i)}(\text{day}_i) \\ \text{P3: } \eta_i &= \beta_0 + f_{a(i)}^*(\text{day}_i) \\ \text{P4: } \eta_i &= \beta_0 + f_1(\text{day}_i) + f_{a(i)}^*(\text{day}_i) \end{aligned}$$

where $a(i)$ indicates to which animal the i th observation belongs. Model P1 includes the mean weight of each animal through a parametric factor term, $\beta_{a(i)}$, plus a smooth of observation day day_i *per* animal. The parametric factor term is required because the animal-specific smooths in this model are each subject to the sum-to-zero-constraint and as such do not contain the constant functions that are needed to model average weight of each animal. These smooths are known informally as “factor by smooths”. Model P2 decomposes the data into an average growth curve, $f_1(\text{day}_i)$, plus smooths, one per animal, that represent deviations from the average smooth, $f_{a(i)}(\text{day}_i)$. Unlike the *factor by smooths*, these deviation smooths do include constant terms for each animal’s average weight, hence only a constant term, β_0 is include in the parametric part of this model. In both models P1 and P2, the $f_{a(i)}(\text{day}_i)$ smooths each have their own smoothing parameters, allowing the wiggleness of each animal’s growth curve to vary, if supported by the data.

Model P3 is similar to model P1, except that each animal’s growth curve, $f_{a(i)}^*(\text{day}_i)$, shares a single smoothing parameter for the wiggleness and hence assumes that the wiggleness of each curve is similar. These smooths are denoted with a superscript $*$ to indicate the shared smoothing parameter, and can be thought of as the smooth equivalent of random slopes and intercepts; informally, we refer to these as *random smooths*.

260 These smooths are fully penalised and contain constant terms to model the average
261 weight of each animal, hence only the intercept, β_0 is included in the parametric part
262 of the model. Model P4 is similar to model P2 in terms of the decomposition into an
263 *average* smooth and animal-specific smooth deviations, but, like model P3, the animal-
264 specific smooths share a smoothing parameter and therefore assume they have similar
265 wiggliness.

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

269 *Japanese quail*

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

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

280 The quail weight data were provisionally assumed to be conditionally gamma dis-
281 tributed. Despite this being a reasonable working assumption, model diagnostics
282 identified deviations from this assumption, and instead the data were modelled as
283 being conditionally Tweedie distributed. The Tweedie family of distributions contains
284 the Poisson (power, $p = 1$) and gamma distributions ($p = 2$) as special cases, and
285 is more flexible than gamma, and allows for a range of mean-variance relationships

through the power parameter, p , of the distribution. In the Tweedie GAMs that were fitted, the power parameter p was allowed to vary between 1–2 and was estimated as an additional model constant term. Models fitted were:

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

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

$$\begin{aligned} \text{Q3: } \eta_i = & \beta_0 + f(\text{day}_i) + f_{\text{treat}(i)}(\text{day}_i) + f_{\text{sex}(i)}(\text{day}_i) + f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i) \\ & + f_{\text{egg}(i)}^*(\text{day}_i) + \xi_{\text{mother}(i)} \end{aligned}$$

$$\begin{aligned} \text{Q4: } \eta_i = & \beta_0 + f(\text{day}_i) + f_{\text{treat}(i)}(\text{day}_i) + f_{\text{sex}(i)}(\text{day}_i) + f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i) \\ & + \psi_{\text{egg}(i)} + \xi_{\text{mother}(i)} \end{aligned}$$

where day_i is the number of days since hatching, $\text{treat}(i)$, $\text{sex}(i)$, and $\text{egg}(i)$ indicate to which treatment group, sex, and bird the i th observation belongs. Smooth functions with subscript $\text{treat}(i)$, $\text{sex}(i)$, or $\text{egg}(i)$ are factor-smooth interactions representing deviations from the “average” smooth, $f(\text{day}_i)$, for the indicated factor. $f_{\text{treat}(i), \text{sex}(i)}(\text{day}_i)$ represents a higher order factor-smooth interaction, which allows the time varying treatment effects to also vary between male or female quail. $\psi_{\text{egg}(i)}$ and $\xi_{\text{mother}(i)}$ are iid Gaussian random intercepts for individual quail and their mother respectively. A f^* represents a random smooth, where each smooth in the set shares a smoothing parameter. The factor-smooth interactions in this model were fitted using the constrained factor-smooth interaction basis in the *mgcv* package; this basis excludes the main effects (and lower-order terms in the case of higher-order interactions) to insure that the basis functions are orthogonal to those lower order terms.

Model Q1 represents a null model containing no treatment effects but models the remaining features of the data, decomposing the growth curves into an average effect, a

sex-specific effect, and individual quail-specific effects, plus a maternal effect. Model Q2 extends Q1 by adding the treatment effect, $f_{\text{treat}(i)}(\text{day}_i)$, which models deviations from the average curve for each of the four treatment levels. Model Q3 further extends the Q2 to allow different treatment-specific curves for male and female quail. Model Q4 is a variant of Q3, replacing the quail-specific growth curve deviations with an individual level random intercept. *A priori*, model Q3 represents the complete set of hypotheses an analyst might expect to consider for these data.

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

Smoothness selection & inference

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

In the lactation curve example I use standard model appraisal techniques to identify the most useful model. In the pig growth example, AIC was used to identify which of the decompositions of time provided the best fit, but the specific choice of model was made on other grounds, using domain knowledge.

In the quail hormone example, AIC values are reported, but *a priori* I chose to report

335 results from model Q3, the “full” model from the point of view of the potential hypotheses
336 under consideration; as the treatment effect was allowed to vary between sexes, I take
337 an estimation-based approach and quantify the magnitudes of any treatment by sex in-
338 teractions using the most complex model. The remaining models are fitted and reported
339 on briefly for illustrative purposes; performing model term selection (for example, testing
340 the higher order $f_{\text{treat}(i)}(\text{day}_i)$ interaction, and deciding on the basis of a p value or AIC
341 whether to retain it in the model or not), would introduce biases into the inference pro-
342 cess. As we currently lack good post selection inference methods for handling this kind
343 of selection, it is prudent to simply not enter into such selection procedures.

344 Using model Q3 allows there to be treatment differences between male and female
345 quail; subsequent estimation of values of interest and comparison among these es-
346 timates is instead the inference route followed for the quail example. To illustrate, I
347 estimate two quantities of interest: i) the estimated growth rate (slope) of an average
348 quail at $\text{day} = 20$ in all combinations of treatment and sex, and ii) the estimated weight
349 of an average quail at the end of the experiment ($\text{day} = 78$), again for all combina-
350 tion of treatment and sex. All pairwise comparisons among treatment levels within sex
351 were performed, with adjustment of p values to control the false discovery rate using the
352 Benjamini–Yekutieli ([Benjamini and Yekutieli, 2001](#)) procedure. Estimates and pairwise
353 comparisons were computed using the `slopes()` and `predictions()` functions of the
354 *marginaleffects* package for R (version 0.27.0, [Arel-Bundock et al., 2024](#)). Estimates
355 of the expected growth curves for average quail in all combinations of treatment and
356 sex were produced using the `conditional_values()` function in the R package *gratia*
357 (version 0.10.0.9018, [Simpson, 2024](#)).

358 All figures were produced using the R packages *gratia* and *ggplot2* (version 3.5.2, [Wick-](#)
359 [ham, 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) and the Tweedie GLM (Figure 3c) show a significant amount of unmodelled signal, while the response residuals for the Tweedie GAM (Figure 3d) are much smaller and do not show a residual pattern. Despite using roughly twice as many degrees of freedom as the other models (Table 1), the Tweedie GAM was clearly favoured in terms of AIC and root mean squared error of the fitted values.

Pig growth

The estimated growth curves for the 18 pigs in the pig growth example are shown in Figure 4, which were produced from model P2. Of the four models fitted, the two models that decomposed the growth effects into an average curve plus animal-specific deviations from the average curve, models P2 and P4, resulted in the most parsimonious fits from the point of view of AIC Table 2. This is due to these two model forms using many fewer degrees of freedom (~64) than either of models P1 (EDF = 83.54) or P3 (EDF = 90.57). Models P1 and P3 do not include the average curve and as a result expend many more degrees of freedom modelling the same general shape for each animal. Interestingly, for these data at least, allowing for a different smoothing parameter (P2) for each animals' deviation smooth seems to be preferred over using a single smoothing parameter (P4). In part, this is due to the somewhat idiosyncratic nature of each animal's

387 growth curve in this data set.

388 Although the deviance explained is very high (~96%) for all models, much of this is
389 due to use of random effects to model differences between animals, and should not
390 be taken as a sign that the model can effectively perfectly predict the weight of a pig
391 under similar conditions; it is clear from Figure 4 that there is much unmodelled variation
392 around the estimated growth curves. One feature of the data that I do not address here
393 is the clear variation among animals in the variance of the depth camera-based weight
394 measurements; animals 5, 6, and 9 in particular, exhibit substantially greater variation
395 than the other animals in the data set.

396 Table 3 provides an overview of the two terms in model P2. Although the average curve
397 ($f(\text{day}_i)$) was allowed to use $k = 9$ basis functions, the wiggleness penalty has shrunk
398 this back to 7.332 effective degrees of freedom (EDF). The animal-specific deviation
399 smooths, $f_{a(i)}(\text{day}_i)$, were fully penalised and as such used $k = 10$ basis functions per
400 animal. Here we clearly see the effect of the wiggleness penalty, which has resulted in
401 a reduction from a potential 180 EDF for the set of animal-specific smooths to 56.054
402 EDF. The test of the null hypothesis for the average growth curve and the omnibus test
403 for the pig-specific deviation smooths indicate both are statistically interesting.

404 *Quail hormone experiment*

405 As assessed by AIC, models containing treatment effects resulted in no improvement
406 in the model fit (Table 4), a result that is consistent with the findings of Sarraude et al.
407 (2020a). However, the results of model Q3 are reported below as this model is con-
408 sistent with the *a priori* assumption of treatment effects, possibly varying by sex, that
409 underlay the original experiment. Using AIC to perform model selection would invalidate
410 subsequent statistical inference we might wish to conduct on the fitted model. Figure 5
411 shows the original data and the estimated growth curves for each individual quail that
412 were obtained from model Q3. The need for quail-specific random smooths is clear;
413 model Q4 had the same model structure as that of Q3, except for the replacement of

414 the quail-specific random smooths with quail-specific intercepts, which resulted in an
415 increase in AIC of over 500 units (Table 4).

416 To focus on the estimated treatment effects, model Q3 was evaluated at 100 evenly-
417 spaced values over the time covariate (the number of evaluation points chosen to ob-
418 tain a visually smooth representation of the estimated functions), plus all combinations
419 of treatment level and sex. The effects of the quail-specific random smooths and the
420 maternal random effects were excluded from these estimates. This results in estimated
421 treatment effects for the average quail, which, strictly speaking should not be interpreted
422 as population level effects due to the non-identity link function. The resulting estimated
423 effects are shown in Figure 6. While there are clear differences in the growth curves
424 based on sex, there appears to be little qualitative difference in the effects of treatment
425 levels on quail growth.

426 The statistical summary of model Q3 is shown in Table 5. The omnibus tests of the treat-
427 ment specific deviations, and the treatment by sex deviations from the average curve,
428 both have $p > 0.05$, which further reinforces the previously made observations that the
429 effects of exposure to maternal thyroid hormone, if any, are small relative to the uncer-
430 tainty in the model itself. The p value for the omnibus test of the quail-specific deviation
431 smooths ($f_{\text{egg}(i)}(\text{day}_i)$) is also greater than 0.05. This is somewhat surprising, given the
432 magnitude of the difference in AIC that is observed when these quail-specific random
433 smooths are replace by a simple random intercept term ($\Delta\text{AIC} = 525.548$). Despite
434 their lacking “statistical significance”, removing these terms on the basis of p values
435 would lead to invalid inference.

436 The largest contribution to the overall EDF of this model (EDF = 274.43) is from the
437 quail-specific deviation smooths (EDF = 221.359), although as there are 57 individual
438 animals and the random smooths include random intercepts to account for differences
439 between the average weight of each quail, this is unsurprising. Again, we see the effect
440 of the wiggleness penalty, which has penalised the smooths back to EDF = 221.359 from

441 a maximum of EDF of 342.

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

450 Finally, pairwise comparisons of quail weight at the end of the experiment (day 78)
451 among treatment levels within sex were used to examine differences in estimated weight
452 of quail due to exposure to maternal thyroid hormones. The results of these comparisons
453 are shown in Table 7. The largest observed difference (~10.1 g) in weight of an aver-
454 age quail was between the T_4 hormone treatment and controls in female quail. This
455 represents approximately a 5% decrease in the weight of an average female quail when
456 exposed to the T_4 hormone compared with the untreated controls. Yet, given the uncer-
457 tainty in the estimates of the model coefficients, the 95% credible interval includes 0 for
458 this, and all other, comparisons.

459 **Author' Points of View**

460 The results presented above, clearly demonstrate the utility of GAMs for modelling
461 the kinds of data commonly encountered in research involving animals. GAMs are of-
462 ten viewed unfavourably as being subjective, requiring the user to specify how complex
463 they want the estimated smooth functions to be, data hungry, and more of a data visual-
464 isation tool ill-suited to formal statistical inference. These views, while perhaps valid in
465 the case of the subjectivity critique, are largely out-dated with modern GAMs as I have
466 presented above. The subjectivity critique has been addressed through developments

467 in automated smoothness selection. None of the data sets analysed here are especially
468 large, certainly not given today's standards, and much research has been done (e.g., [Li](#)
469 [and Wood, 2020](#)) and continues to be done to adapt the algorithms to ever higher di-
470 mensional problems. The `bam()` function in the *mgcv* package, for example, can handle
471 data on the order of millions of rows, with tens of thousands of model parameters, given
472 sufficient availability of computer memory.

473 As demonstrated in the quail maternal hormone example, formal statistical inference
474 is entirely feasible. If we ignore the selection of smoothness parameters, the models
475 described here are little more than generalized linear (mixed) models (GL(M)Ms) once
476 the basis expansions of covariates have been performed. Modern software tools like the
477 *marginaleffects* package for R allow a consistent interface to statistical inference across
478 many disparate statistical models, and GAMs are no different in this regard.

479 The distinct advantage of GAMs over GL(M)Ms is their ability to learn the shapes of
480 nonlinear relationships between covariates and the response from the data themselves.
481 This relieves the analyst from having to force data to fit a particular theoretical model,
482 unless they have a good justification to use that model of course. GAMs also avoid
483 the model selection problems inherent to modelling with polynomial basis expansions
484 (e.g. $x + x^2 + x^3 + \dots + x^p$).

485 The main disadvantage of GAMs is that they are more complex to fit than GL(M)Ms;
486 the analyst has some additional data modelling choices to make when using GAMs. The
487 main choice is the number of basis functions, K , that should be used by each smooth
488 in the model. The general advice here is for the analyst to imagine the largest amount
489 of wiggleness that they would expect and then set k a little larger than this. Of course,
490 the novice user of GAMs will lack the experience required to do this easily, but for the
491 sorts of data problems exemplified above, we would not expect highly complex smooth
492 functions, and the default of $k = 10$ for univariate smooths in *mgcv* is usually sufficient
493 for many problems. The key requirement is that the initial number of basis functions

494 used should be large enough such that the span of functions representable with that
495 basis will contain the true but unknown function or a close approximation to it. The basis
496 dimension must be checked of course, and this adds another step to the model appraisal
497 or checking procedure. With *mgcv* for example, the `k.check()` function provides a test
498 for sufficiency of the basis size used to fit the model (Pya and Wood, 2016).

499 A further disadvantage is that statistical inference is somewhat more approximate than
500 with GL(M)Ms. While GAMs share with GL(M)Ms the property of having asymptotically
501 correct p values, the p values for smooths are more approximate than for terms in a
502 GL(M)M because the current theory on which these tests are based does not account
503 for the selection of smoothing parameters; for the purposes of the tests, the smoothing
504 parameters are treated as being fixed and known, but instead they are estimated from
505 the data (Wood, 2013). While there has been some progress in adapting the theory to
506 include this additional source of uncertainty (e.g., Wood et al., 2016), as yet this has not
507 been used to correct the p values of tests of smooths.

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

515 GAMs are a very broad and general class of models. In the case of the pig growth data,
516 several animals exhibited visibly more variation in their weight measurements than the
517 majority of the animals in the data. Such heteroscedasticity (non-constant variance)
518 can be modelled using distributional GAMs (or Generalized additive models for location,
519 scale, shape or GAMLSS) (e.g. Rigby and Stasinopoulos, 2005; Kneib et al., 2021; Klein,
520 2024), which include linear predictors for all of the parameters of a distribution, or centile

521 or quantile models (e.g., [Nakamura et al., 2022](#)). In the examples above, the additional
522 parameters (scale in the case of the Gamma, or the Tweedie power parameter) were
523 estimated as constants for the entire data set. A distributional GAM would allow those
524 parameters to potentially vary with individual animals, or as as smooth functions of the
525 covariates, just as was done for the mean of the distribution in this study.

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

529 **Ethics approval**

530 Not applicable. See the original sources of the data used for ethical approvement.

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

532 The author did not use any artifical intelligence technologies in the writing process.

533 **Author ORCIDs**

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

535 **Author contributions**

536 GLS: Conceptualization, Methodology, Software, Formal analysis, Writing, Visulaiza-
537 tion.

538 **Declaration of interest**

539 None.

540 Acknowledgements

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

546 Financial support statement

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

549 References

- 550 Arel-Bundock, V., Greifer, N., Heiss, A., 2024. How to interpret statistical models using marginaeffects for
551 *R* and *Python*. J. Stat. Softw. 111, 1–32. URL: <https://www.jstatsoft.org/index.php/jss/article/view/v111i09>,
552 doi:10.18637/jss.v111.i09.
- 553 Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency.
554 Ann. Stat. 29, 1165–1188. URL: http://projecteuclid.org/download/pdf_1/euclid.aos/1013699998.
- 555 Benni, S., Pastell, M., Bonora, F., Tassinari, P., Torreggiani, D., 2020. A generalised additive model to
556 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)
557 [S1751731119001721](http://dx.doi.org/10.1017/S1751731119001721), doi:10.1017/S1751731119001721.
- 558 Brito, L.F., Gomes da Silva, F., Rojas de Oliveira, H., Souza, N., Caetano, G., Costa, E.V., Romeiro de
559 Oliveira Menezes, G., Puerro de Melo, A.L., Teixeira Rodrigues, M., de Almeida Torres, R., 2017. Modelling
560 lactation curves of dairy goats by fitting random regression models using legendre polynomials or b-splines.
561 Can. J. Anim. Sci. URL: <http://dx.doi.org/10.1139/cjas-2017-0019>, doi:10.1139/cjas-2017-0019.
- 562 Bus, J.D., Franchi, G.A., Boumans, I.J.M.M., Te Beest, D.E., Webb, L.E., Jensen, M.B., Pedersen, L.J.,
563 Bokkers, E.A.M., 2025. Short-term associations between ambient ammonia concentrations and growing-
564 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)
565 [prevetmed.2025.106555](http://dx.doi.org/10.1016/j.prevetmed.2025.106555), doi:10.1016/j.prevetmed.2025.106555.
- 566 Franchi, G.A., Bus, J.D., Boumans, I.J.M.M., Bokkers, E.A.M., Jensen, M.B., Pedersen, L.J., 2023. Estimating
567 body weight in conventional growing pigs using a depth camera. Smart Agric. Technol. 3, 100117. URL:
568 <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.
- Nagel-Alne, G.E., Krontveit, 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.

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](https://arxiv.org/abs/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.

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:

639 <http://dx.doi.org/10.1111/1467-9868.00374>, doi:10.1111/1467-9868.00374.

640 Wood, S.N., 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semi-
 641 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.

642 <http://dx.doi.org/10.1111/j.1467-9868.2010.00749.x>, doi:10.1111/j.1467-9868.2010.00749.x.

643 Wood, S.N., 2013. On p-values for smooth components of an extended generalized additive model. Biometrika
 644 100, 221–228. URL: <http://biomet.oxfordjournals.org/content/100/1/221.abstract>, doi:10.1093/biomet/
 645 ass048.

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

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

652 Yano, M., Shimadzu, H., Endo, T., 2014. Modelling temperature effects on milk production: a study on holstein
 653 cows at a japanese farm. SpringerPlus 3, 129. URL: <http://dx.doi.org/10.1186/2193-1801-3-129>, doi:10.
 654 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 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 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	3823.751	2.050	0.964
P4	64.754	3809.395	2.148	0.963

Table 3: Model 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, F is the test statistic and p the p value of the null hypothesis of a flat constant function or functions.

Model term	Label	k	EDF	F	p
$f(\text{day})$	s(day)	9	7.332	1302.500	<0.001
$f_{a(i)}(\text{day})$	s(day,animal)	10	56.054	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 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	4458.922	12.360	0.998
Q3	274.430	4461.062	12.455	0.998
Q4	77.115	4986.609	30.416	0.995

Table 5: Model 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, F is the test statistic and p the p value of the null hypothesis of a flat constant function or functions.

Model term	Label	k	EDF	F	p
$f(\text{day})$	s(day)	9	8.947	4785.756	<0.001
$f_{\text{treat}(i)}(\text{day}_i)$	s(day,treat)	10	8.307	1.501	0.1467
$f_{\text{sex}(i)}(\text{day}_i)$	s(day,sex)	10	7.322	14.393	<0.001
$f_{\text{treat}(i),\text{sex}(i)}(\text{day}_i)$	s(day,treat,sex)	10	9.807	1.596	0.0984
$f_{\text{egg}(i)}(\text{day}_i)$	s(day,egg)	6	221.359	67.155	0.2666
$\zeta_{\text{mother}(i)}$	s(mother)	21	17.688	9.055	<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	$T_3 - \text{Control}$	-0.339	0.290	-1.168	1	-1.148	0.471
Female	$T_3 T_4 - \text{Control}$	-0.395	0.298	-1.326	1	-1.227	0.437
Female	$T_4 - \text{Control}$	-0.495	0.356	-1.389	1	-1.490	0.500
Female	$T_3 T_4 - T_3$	-0.056	0.218	-0.259	1	-0.666	0.553
Female	$T_4 - T_3$	-0.156	0.231	-0.676	1	-0.802	0.490
Female	$T_4 - T_3 T_4$	-0.100	0.275	-0.364	1	-0.866	0.667
Male	$T_3 - \text{Control}$	-0.194	0.299	-0.649	1	-1.031	0.642
Male	$T_3 T_4 - \text{Control}$	-0.251	0.263	-0.957	1	-0.984	0.482
Male	$T_4 - \text{Control}$	-0.388	0.339	-1.144	1	-1.335	0.559
Male	$T_3 T_4 - T_3$	-0.057	0.262	-0.217	1	-0.788	0.674
Male	$T_4 - T_3$	-0.194	0.273	-0.708	1	-0.957	0.570
Male	$T_4 - T_3 T_4$	-0.137	0.294	-0.465	1	-0.957	0.684

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	$T_3 - \text{Control}$	-6.332	5.838	-1.085	1	-22.682	10.017
Female	$T_3 T_4 - \text{Control}$	-5.577	5.614	-0.993	1	-21.298	10.144
Female	$T_4 - \text{Control}$	-10.087	5.762	-1.751	1	-26.224	6.050
Female	$T_3 T_4 - T_3$	0.756	4.344	0.174	1	-11.411	12.923
Female	$T_4 - T_3$	-3.755	4.038	-0.930	1	-15.064	7.555
Female	$T_4 - T_3 T_4$	-4.510	3.648	-1.236	1	-14.727	5.707
Male	$T_3 - \text{Control}$	2.493	5.391	0.462	1	-12.605	17.592
Male	$T_3 T_4 - \text{Control}$	-2.992	4.035	-0.742	1	-14.292	8.308
Male	$T_4 - \text{Control}$	-0.808	4.579	-0.176	1	-13.632	12.017
Male	$T_3 T_4 - T_3$	-5.485	4.811	-1.140	1	-18.960	7.990
Male	$T_4 - T_3$	-3.301	4.709	-0.701	1	-16.489	9.887
Male	$T_4 - T_3 T_4$	2.184	3.740	0.584	1	-8.291	12.659

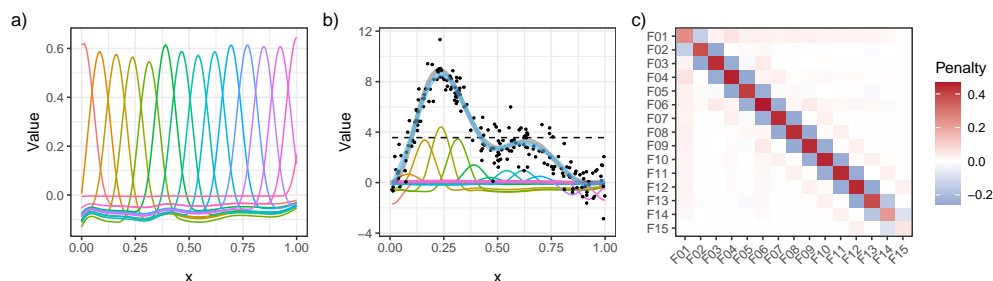


Fig. 1: Illustration of how penalised splines work. A spline basis expansion (a) and associated penalty matrix S (c) are formed for a covariate x . Model fitting involves finding estimates for the coefficients of the basis functions that make the fitted spline (thick, blue curve) go as close to the data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

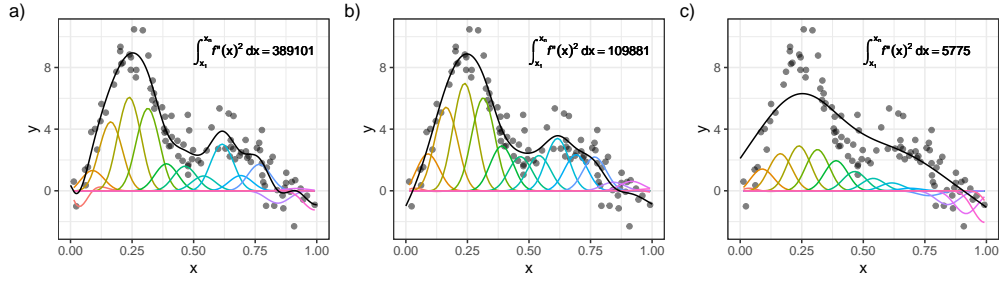


Fig. 2: Illustration of how the wigginess penalty controls the resulting fit of a penalised spline. The weighted basis functions are shown as thin coloured lines. In each panel a penalised spline is shown by the solid black line, which has been fitted to the data points shown. The wigginess value of 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, λ , is used as a tuning parameter in the model, which ultimately controls this balance between fit and complexity.

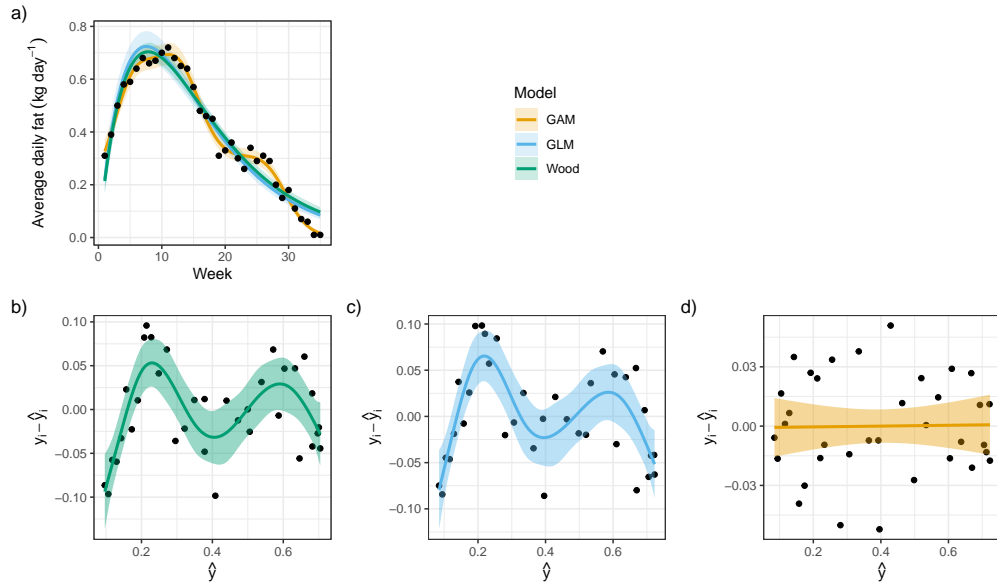


Fig. 3: Results of model fitting to the average daily fat content data from Henderson and McCulloch (1990). a) observed average daily fat content (points) and estimated lactation curves from Wood's (1967) model, a Tweedie GLM, and a Tweedie GAM (lines) with associated 95% confidence (Wood's model) or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

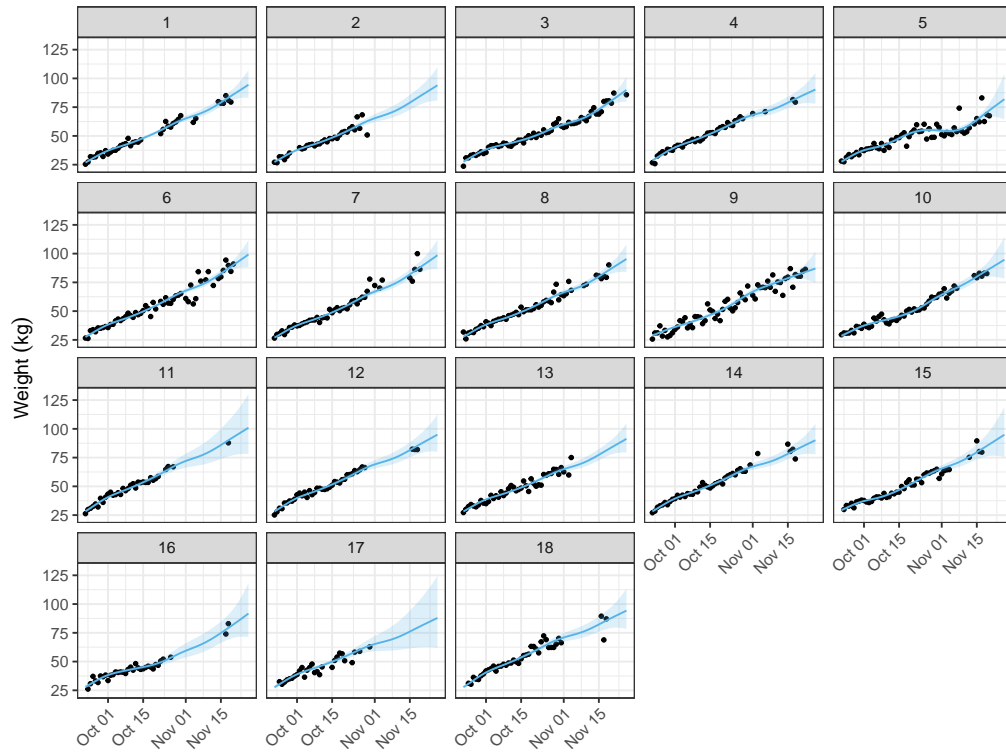


Fig. 4: Depth camera-based weight estimates from 18 commercial pigs. The data (black points) are the average of multiple measurements taken of each animal per day, 1 panel per pig. The panel labels indicate to the pig shown. The estimated growth curve for each pig obtained using generalized additive 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.

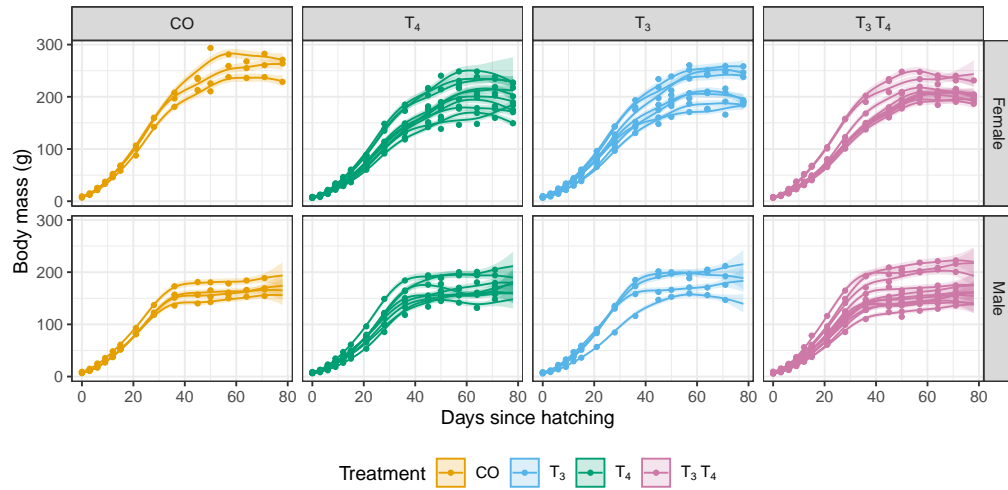


Fig. 5: Measured body mass (g) for 57 Japanese quail from an experiment on the effects of exposure to the maternal hormone T₄, its active metabolite T₃, both (T₃T₄) or a saline solution control (CO). The data (points) are shown along side the estimated growth curve for each quail obtained using generalized 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.

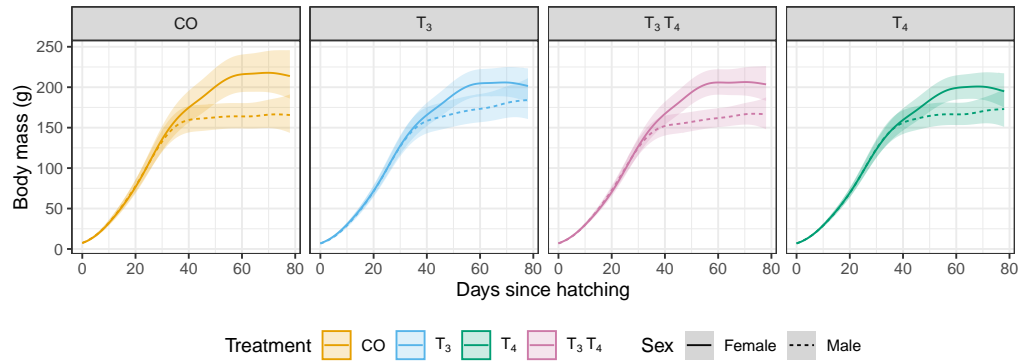


Fig. 6: Conditional value plots showing the growth rate for an average quail in each of the treatment groups by sex. The panels show the estimated curves for a particular treatment group; saline solution controls (CO), maternal hormone metabolite T₃, the maternal hormone T₄, and a combination of both T₃T₄. 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.

655 Figure captions

656 Fig. 1: Illustration of how penalised splines work. A spline basis expansion (a) and
 657 associated penalty matrix S (c) are formed for a covariate x . Model fitting involves finding
 658 estimates for the coefficients of the basis functions that make the fitted spline (thick, blue
 659 curve) go as close to the data (black points) as possible, without over fitting (b). In (a)
 660 and (b) the basis functions are shown as thin coloured lines and are from a B spline basis.
 661 The sum-to-zero identifiability constraint needed so that an intercept can be included in
 662 the model has been absorbed into the basis shown. The dashed horizontal line in (b) is
 663 the estimated value of the intercept. The penalty matrix (c) encodes how wiggly each
 664 basis function is in terms of its second derivative.

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

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

Fig. 4: 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 to 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. 5: Measured body mass (g) for 57 Japanese quail from an experiment on the effects of exposure to the maternal hormone T_4 , its active metabolite T_3 , both (T_3T_4) 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. 6: 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.