**Supplementary Material**

**Toward better estimates of the real-time individual amino acid requirements of growing-finishing pigs showing deviations from their typical feeding patterns**

L. Hauschild 1, A. R. Kristensen 2, I. Andretta 3, A. Remus 4, L. S. Santos 1 and C. Pomar 4

Animal Journal

**S1 – Posterior and prior distribution – Kalman Filter**

The mean and variance-covariance matrix of the posterior distribution are represented as **m*t*** and **C*t***, respectively, such that (θ*t*|D*t*) ~ *N*(m*t*, C*t*). In order to initialize the model, prior information is required (**D0**; before any observations are made, *t* = 0) on the initial distribution of the parameter vector: (θ0|D0) ~ *N*(m0, C0). The observation and evolution error sequences *vt* and *wt* are assumed to be internally and mutually independent and are independent of (θ0|D0).

The recursively obtained prior distribution for θ*t* at time *t* − 1 for the Kalman filter is described as follow:

$\left(D\_{0}\right)\~N\left(a\_{t}, R\_{t}\right)$ [S1]

where **a*t*** = G*t*m*t*−1 and **R*t***= G*t*C*t*−1G′*t* + W*t*. The one-step forecast for Y*t* at time *t* is as follows:

$\left(D\_{t-1}\right)\~N\left(f\_{t}, Q\_{t}\right)$ [S2]

where **f*t*** = F′*t*a*t* and **Q*t***= F′*t*R*t*F*t* + V*t*. Finally, the posterior distribution for θ*t* at time *t* is as follows:

$\left(D\_{t}\right)\~N\left(m\_{t}, C\_{t}\right)$ [S3]

where m*t* = a*t* + A*t*e*t* and C*t* = R*t* − A*t*Q*t*A*t*′, with the adoptive matrix (**A*t***) specified as A*t* = R*t*F*t*Q*t*−1. The vector of one-step forecast errors (**e*t***) is calculated as e*t* = Y*t* − f*t*. The vector m*t* and the matrix C*t* are referred to as the filtered mean and variance-covariance matrix of the parameter vector at time *t*, respectively.

Sequential forecast for k steps ahead is calculated as follows for *j* = 1, …, *k*:

$\left(D\_{t}\right)\~N\left[a\_{t}\left(j\right), R\_{t}\left(j\right)\right] $ [S4]

where a*t*(*j*) = G*t+j*a*t*(*j* − 1) and R*t*(*j*) = G*t+1*R*t*(*j* − 1)G*t+1* + W*t+j*, with the initial values a*t*(0) = m*t* and R*t*(0) = C*t*. Based on this parameter vector distribution, the following forecast distribution is obtained:

$\left(D\_{t}\right)\~N\left[f\_{t}\left(j\right), Q\_{t}\left(j\right)\right] $ [S5]

where f*t*(*j*) = F′*t+j*a*t*(*j*) and Q*t*(*j*) = F′*t+j*R*t*(*j*)F*t+j* + V*t+j*.

The proposed model works with one-step-ahead forecast (*j* = 1).

Based on retrospective analysis, moments for θ*t−j* given all observations D*t* are specified as follows:

$\left(D\_{t}\right)\~N\left[a\_{t}\left(-j\right), R\_{t}\left(-j\right)\right] $ [S6]

where a*t*(−*j*) = m*t−j* + B*t−j*[a*t*(*−j* + 1) − a*t−j*+1] and R*t*(*−j*) = C*t−j* + B*t−j*[R*t*(−*j* + 1) − R*t−j*+1]B*t−j*, with B*t−j* = C*t−j*G′*t−j+1*R−1*t−j+1* and given initial values for a*t*(0) = m*t* and R*t*(0) = C*t*. The vector a*t*(−*j*) and the matrix R*t*(−*j*) are also denoted as m̃*t−j* and C̃*t−j*, respectively, and are referred to as the smoothened mean and variance-covariance matrix at time *t* − 1, respectively.

**S2 - Cumulative tabular sum**

This method accumulates deviations from **T0** (target value) that are above the target with one statistic, **C+**, and those that are below the target with another statistic, **C−**. The C+ and C− for a given day (*t*) were as follows:

 $C\_{t}^{+}= max\left\{0, e\_{t}^{norm}-\left(T\_{0}+K\right)+C\_{t-1}^{+}\right\}$ [S7]

 $C\_{t}^{-}= min\left\{0, \left(T\_{0}-K\right) e\_{t}^{norm}+C\_{t-1}^{-}\right\}$ [S8]

where T0 = 0, and **K** is the reference value expressed as K = (1 \* $σ\_{t}$)/2. Alarms are activated if $C\_{t}^{+}$ or $C\_{t}^{-}$ exceed a threshold **H** (expressed in terms of the SD) in a given day *t*. The Cumulative tabular sum (CUMSUM) parameters (the reference value [K] and the decision interval [H]) were selected in order to provide good average run length performance. The K is relevant to the size of a shift that we want to detect, while the H determines in-control and out-of-control performance. If K = 1, we aim to detect a two-SD shift. The choice of which value of K and which value of H should be applied depends on how sensitive the farm manager thinks that the system should be. The starting values of $C\_{t}^{+}$ or $C\_{t}^{-}$ are set to zero. As we are interested only in identifying reductions in feed intake, only the alarms generated when $C\_{t}^{-} $exceeds the threshold H are considered.